

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

3M nucleic - nucleic search, using sw model
Run on: March 10, 2004, 08:38:17 ; Search time 3862.62 Seconds
(without alignments)
17011.249 Million cell updates/sec

Title: US-10-084-817-8
Perfect score: 1516
Sequence: 1 cttgttttggacatagct.....gtgaaataaaatgncgtc 1516

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
GenEmbl:
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_scs.*
28: em_un.*
29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rtd.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1389.4	91.6	1399	6	AX440468 Sequence
2	1387.8	91.5	1399	6	AX380797 Sequence
3	1387.8	91.5	1399	6	AX004164 Sequence
4	1387.8	91.5	1399	6	AX281593 Sequence
5	1387.8	91.5	1399	6	AX334637 Sequence
6	1387.8	91.5	1399	6	AX779855 Sequence
7	1387.8	91.5	1399	9	AX05908 Homo sapien
8	1375.8	90.8	1398	6	AX004167 Sequence
9	1367.4	90.2	1408	9	BC001275 Homo sapi
10	1364.8	90.0	1376	6	I02238 Sequence 1
11	1364.8	90.0	1376	6	I02432 Sequence 3
12	1364.8	90.0	1376	6	I08414 Sequence 20
13	1357.8	89.6	1385	9	BC035993 Homo sapi
14	1329.8	87.7	2239	9	AK074480 Homo sapi
15	976	64.4	1252	4	AF544227 Equus cab
16	947.6	62.5	1342	4	AF544227 Equus cab
17	938.4	61.9	1335	4	SSANEXNI
18	917.8	60.5	1327	4	BIANNI
19	913.8	60.3	1379	10	CCLIIPP33
20	886.8	58.5	1402	6	AX401909
21	886.8	58.5	1402	10	NMLCI
22	885.2	58.4	1355	10	NMLCI
23	884.8	58.4	1409	10	BC061710
24	882.2	58.2	1383	10	BC002289
25	880	58.0	1353	6	RATCALPII
26	880	58.0	1355	6	E01751
27	879.4	58.0	1933	6	E01979
28	877.2	57.9	1358	10	BC004594
29	852	56.2	1330	10	MUSLCI
30	599.4	38.5	1369	5	PGNANNI
31	589.4	38.9	614	6	AX440388
32	566	37.3	567	6	AX885063
33	566	37.3	567	6	BD024673
34	555.4	36.6	1224	5	PGNCAL
35	540.2	35.6	1237	5	BC053786
36	520.4	34.3	550	6	AX197684
37	519.4	34.3	537	6	AX208294
38	457.6	30.2	1409	5	BC064261
39	453.2	29.9	1235	5	AX178793
40	451.6	29.8	1206	5	BC053190
41	446.2	29.4	1337	5	BC054187
42	424.4	28.0	438	6	AX321079
43	422.8	27.9	1544	5	OLANNEXM3
44	362.8	23.9	542	6	BD216773
45	349.6	23.1	1280	5	BC041306

ALIGNMENTS

RESULT 1
AX440468
LOCUS AX440468 1399 bp DNA linear PAT 28-JUN-2002
DEFINITION Sequence 321 from Patent WO0190154.
ACCESSION AX440468
VERSION AX440468.1 GI:21665278
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Xu, J., Mitcham, J.L., Harlocker, S.L., Dillon, D.C., Secret, H.,
Lodes, M.J., Algate, P.A., Fling, S.P., Mannion, J., Benson, D.R. and
Carter, D.

QY	DB	841	TGAAGGTGACATTTGAGAAATGCTTCACAGCTATCTGCTGAAGTGGGCCCAAGCAACACG	900
1025	QY	1025	CTTTCTTTTGAGAGAAAGCTTCATCAAGCCATGAAGAGTGTGGAACTGCTCCCATAGGCAT	1084
901	DB	901	CTTTCTTTTGAGAGAAAGCTTCATCAAGCCATGAAGAGTGTGGAACTGCTCCCATAGGCAT	960
1085	QY	1085	TGATCAGGATATATGGTTTCCCGTTCCTGAAATTCATCAAGAGTGTGGAACTGCTCCCATAGGCAT	1144
961	DB	961	TGATCAGGATATATGGTTTCCCGTTCCTGAAATTCATCAAGAGTGTGGAACTGCTCCCATAGGCAT	1020
1145	QY	1145	AGAAGATGATATGATATCTCCCTTTGGCAAGCCATCCTGGATGAAACCAAGAGGAGTATG	1204
1021	DB	1021	AGAAGATGATATGATATCTCCCTTTGGCAAGCCATCCTGGATGAAACCAAGAGGAGTATG	1080
1205	QY	1205	AGAAAATCCCTGGTGGCTCTTTTGGAGGAGAACTAAACATTCCTTTGATGGTCTCAAGCTA	1264
1081	DB	1081	AGAAAATCCCTGGTGGCTCTTTTGGAGGAGAACTAAACATTCCTTTGATGGTCTCAAGCTA	1140
1265	QY	1265	TGATCAGAGAGCTTTAAATATATATTTTCATCCTATAGCTTAAATAGGAAAGTTTCTTC	1324
1141	DB	1141	TGATCAGAGAGCTTTAAATATATATTTTCATCCTATAGCTTAAATAGGAAAGTTTCTTC	1200
1325	QY	1325	AACAGGATTACAGTGTAGCTACCTACATCTGAAAATATAGCCTTTAAATCATTTTTAT	1384
1201	DB	1201	AACAGGATTACAGTGTAGCTACCTACATCTGAAAATATAGCCTTTAAATCATTTTTAT	1260
1385	QY	1385	ATTATRACTCTGTATATATAGATATAGTCCATTTTTTAAATAATGTTTTCCCAACCATTA	1444
1261	DB	1261	ATTATRACTCTGTATATATAGATATAGTCCATTTTTTAAATAATGTTTTCCCAACCATTA	1320
1445	QY	1445	AAACCCCTATACAAAGTTGTTCTTACTATACATATACATAGAGAAAGATGTCTATGTAGCTGAAAA	1504
1321	DB	1321	AAACCCCTATACAAAGTTGTTCTTACTATACATATACATAGAGAAAGATGTCTATGTAGCTGAAAA	1380
1505	QY	1505	TAAATATGCGTC 1516	
1381	DB	1381	TAAATATGCGTC 1392	
RESULT 2				
AR380797				
LOCUS				
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
FEATURES				
source				
ORIGIN				
Query Match				
Best Local Similarity				
Matches 1389; Conservative				
91.5%; Score 1387.8; DB 6; Length 1399;				
Pred No. 0; Mismatches 3; Indels 0; Gaps 0				
125 AGTGTGAAATCTTCAGAGAAAGATTTCTTTTATGTTCTTTTGCAGAGAGGTAGAGATAAAG				
185 ACACATTTTTTCAAAATATGCAATGTTATCAGAAATCTCTCAAGCAGGCTCGTTTATTGAAA				
61 ACATTTTTCAAAATATGCAATGTTATCAGAAATCTCTCAAGCAGGCTCGTTTATTGAAA				
245 ATGAAGAGCAGGAATATGTTTCAAATCTGTAAGTATCAACCAAGAGTGTTCGGATCAGCGG				

121 ATGAGAGCAGGAATATGTTCAAACTGTGAAGTATCCAAAGGTGGTCCCGATCAGCGG 180
 305 TGAGCCCTATCTCTACCTTCAATCCATCCTCGGATGTCGCTTCCTGCATTAAGGCCATAA 364
 181 TGAGCCCTATCTCTACCTTCAATCCATCCTCGGATGTCGCTTCCTGCATTAAGGCCATAA 240
 365 TGGTTAAGGTTGGATGAAGCAACCATCATGTCATCTTAACCTTAAGGGAACCAATGCAC 424
 241 TGGTTAAGGTTGGATGAAGCAACCATCATGTCATCTTAACCTTAAGGGAACCAATGCAC 300
 425 AGCGTCAACAGATCAAGCAGCATATCTCCAGGAAACAGGAAAGCCCTCGATGAACAC 484
 301 AGCGTCAACAGATCAAGCAGCATATCTCCAGGAAACAGGAAAGCCCTCGATGAACAC 360
 485 TGAAGAAAGCCCTTACAGGTCACTTGAGGAGGTTGTTTACGCTCTGTTAAACCTCCAG 544
 361 TTAAGAAAGCCCTTACAGGTCACTTGAGGAGGTTGTTTACGCTCTGTTAAACCTCCAG 420
 545 CGCAATTTGATGCTGATGAACCTTCTGCTGCTGATGAAGGCTTTGGAACTGATGAAGATA 604
 421 CGCAATTTGATGCTGATGAACCTTCTGCTGCTGATGAAGGCTTTGGAACTGATGAAGATA 480
 605 CTCTAATTTGATGCTGATGAACCTTCTGCTGCTGATGAAGGCTTTGGAACTGATGAAGATA 664
 481 CTCTAATTTGATGCTGATGAACCTTCTGCTGCTGATGAAGGCTTTGGAACTGATGAAGATA 540
 665 ACAGAGAGAACTGAAGAGATCTGGCAAGACATTAACCAAGAAATCAGAGACATTAACAGGCT 724
 541 ACAGAGAGAACTGAAGAGATCTGGCAAGACATTAACCAAGAAATCAGAGACATTAACAGGCT 600
 725 TTCGAAAGCGTTGCTTCTCTGCTGCTGATGAAGGCTTTGGAACTGATGAAGATA 784
 601 TTCGAAAGCGTTGCTTCTCTGCTGCTGATGAAGGCTTTGGAACTGATGAAGATA 660
 785 TAGATCTGCTGATTCAGATGCGAGGCTTGTATGAAGAGGAGAAAGGAGGAGGGA 844
 661 TAGATCTGCTGATTCAGATGCGAGGCTTGTATGAAGAGGAGAAAGGAGGAGGGA 720
 845 CAGAGCTAAACGTTGCTTCAATACCACTTACCAAGAGCTATCCCAACTTCGCGAG 904
 721 CAGAGCTAAACGTTGCTTCAATACCACTTACCAAGAGCTATCCCAACTTCGCGAG 780
 905 TGTTCAGAAATACCAAGTACAGTACGATGATGATGATGATGATGATGATGATGATGATGAT 964
 781 TGTTCAGAAATACCAAGTACAGTACGATGATGATGATGATGATGATGATGATGATGATGAT 840
 965 TGAAGAGTGACATTCAGAAATGCTTACAGCTATCGTGAAGTGGCCCAAGCAACCCAG 1024
 841 TGAAGAGTGACATTCAGAAATGCTTACAGCTATCGTGAAGTGGCCCAAGCAACCCAG 900
 1025 CTTTCTTTGAGAGAGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCT 1084
 901 CTTTCTTTGAGAGAGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCT 960
 1085 TGATCAGGATTTATGTTTCCGCTTCTGAATTTGATGATGATGATGATGATGATGATGATGATGAT 1144
 961 TGATCAGGATTTATGTTTCCGCTTCTGAATTTGATGATGATGATGATGATGATGATGATGATGAT 1020
 1145 AGAAGAT 1204
 1021 AGAAGAT 1080
 1205 AGAAGAT 1264
 1081 AGAAGAT 1140
 1265 TGATCAGAGAGCTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1324
 1141 TGATCAGAGAGCTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1200
 1325 AACAGATACAGTGTAGCTACCTACCTGAGAAATATAGCTTTAATTAATTAATTAATTAATTAAT 1384

1201 AACGAGTTACAGTGTAGCTACCTACCTGCTGAAATATATAGCTTTAAATCATTTTAT 1260
 1385 ATTATACTCTGTATTAATAGAGTAAGTCCATTTTAAATGTTTCCCAACCATTA 1444
 1261 ATTATACTCTGTATTAATAGAGTAAGTCCATTTTAAATGTTTCCCAACCATTA 1320
 1445 AACCCCTATACAGTGTGTTCTAGTAAACATATACATGAGAAAGATGCTATGTAGCTGAAA 1504
 1321 AACCCCTATACAGTGTGTTCTAGTAAACATATACATGAGAAAGATGCTATGTAGCTGAAA 1380
 1505 TAAATGCGTC 1516
 1381 TAAATGCGTC 1392

RESULT 3
 AX004164
 LOCUS
 DEFINITION
 AX004164
 ACCESSION
 AX004164.1
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1
 AUTHORS
 TITLE
 JOURNAL
 GEORGES E. and Wang, Y.
 P-40/annexin i and related proteins and their role in multidrug
 resistance
 Patent: WO 9921980-A 1 06-MAY-1999;
 GEORGES ELIAS (CA); UNIV MCGILL (CA)

FEATURES
 Location/Qualifiers
 1..1399
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 75..1115
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAC04890.1"
 /db_xref="GI:9927715"
 /db_xref="REMBL:CAC04890"
 /translation="MAVSEFLKQAFIENEEQVYVTVKSKGPGSVSPYPTNP
 SDVAALHKAIMVGVDEATIIDLTKNNAQOQIKAAVLOETGKPLDITKKALTG
 HSEVVLALTKTAPQDAPDLAAKMGICGTDEDTLEILASRTNKEIRINRVREL
 KDLADITSDTSGDFRNALLSLAKGDRSEDFVNEADSDARALYBAGERRKQDV
 NVNTILITKSPQLRRVFTQKTKYKSHDMKVLDELKGDIEKCLTAIVKCATSKPA
 FPAEKLHQAMKGVGTHKALIRIMVSRSEIDMNDIKAFYQKMVIGSLCQAILDETQGD
 YEKILVALCGGN"

ORIGIN
 Query Match 91.5%; Score 1387.8; DB 6; Length 1399;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1389; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 125 AGTGTGAATCTTTCAGAGAGAAATTTCTCTTTAGTTCTTTGCAAGAGGTGAGAGATAAG 184
 1 AGTGTGAATCTTTCAGAGAGAAATTTCTCTTTAGTTCTTTGCAAGAGGTGAGAGATAAG 60
 185 ACATCTTTTCAAAAATGGCAATGGTATCAGAAATTCCTCAGCAGGCGCTGTTTATTGAAA 244
 61 ACATCTTTTCAAAAATGGCAATGGTATCAGAAATTCCTCAGCAGGCGCTGTTTATTGAAA 120
 245 ATGAAGAGCAGGAGATATGTTTCAAACTGTGAAGTCATCCAAAGGTGTCGCGATCAGCGG 304
 121 ATGAAGAGCAGGAGATATGTTTCAAACTGTGAAGTCATCCAAAGGTGTCGCGATCAGCGG 180
 305 TGAGCCCTATCTCTACCTTCAATCCATCCTCGGATGTCGCTGCTGCAAGGCCATTA 364
 181 TGAGCCCTATCTCTACCTTCAATCCATCCTCGGATGTCGCTGCTGCAAGGCCATTA 240
 365 TGGTTAAGGTTGGATGAAGCAACCATCATGTCATCTTAACCTTAAGGGAACCAATGCAC 424

Db 241 TGGTTAAAGGTGGTGAAGCAACCATCATTTGACATTTCTAAAGCGAAACAATGCAC 300
 QY 425 AGCGTCAACAGATCAAGACAGCATATCTCCAGAAACAGAAAGCCCTGGATGAACAC 484
 Db 301 AGCGTCAACAGATCAAGACAGCATATCTCCAGAAACAGAAAGCCCTGGATGAACAC 360
 QY 485 TGAAGAAAGCCCTTACAGGTCACTTGAGAGGTTGTTTAGCTGTCTGCTAAACCTCCAG 544
 Db 361 TTAAGAAAGCCCTTACAGGTCACTTGAGAGGTTGTTTAGCTGTCTGCTAAACCTCCAG 420
 QY 545 CGCAATTTGATGCTGATGAATCTGCTGCTCCATGAAGGCTTGGAACTGATGAAGATA 604
 Db 421 CGCAATTTGATGCTGATGAATCTGCTGCTCCATGAAGGCTTGGAACTGATGAAGATA 480
 QY 605 CTCTAATTTGATGCTGATGAATCTGCTGCTCCATGAAGGCTTGGAACTGATGAAGATA 564
 Db 481 CTCTAATTTGATGCTGATGAATCTGCTGCTCCATGAAGGCTTGGAACTGATGAAGATA 540
 QY 665 ACAGAGAGGAACTGAAGAGATCTGCGCAAGACATTAACCTCAGACATCTGGAGATT 724
 Db 541 ACAGAGAGGAACTGAAGAGATCTGCGCAAGACATTAACCTCAGACATCTGGAGATT 600
 QY 725 TTCGAAACGCTTTGCTTTCTCTTCTAAGGCTGACCGATCTGAGGCTTTGGTGTGAATG 784
 Db 601 TTCGAAACGCTTTGCTTTCTCTTCTAAGGCTGACCGATCTGAGGCTTTGGTGTGAATG 660
 QY 785 AAGACTTGGCTGATTCAGATGCCAGGCTTGTATGAGCAGAGAAAGGAGAAAGGGA 844
 Db 661 AAGACTTGGCTGATTCAGATGCCAGGCTTGTATGAGCAGAGAAAGGAGAAAGGGA 720
 QY 845 CAGACGTAACGCTTTCAATACCATCTTACCACAGAGCTATCCACAACTTCGCAGAG 904
 Db 721 CAGACGTAACGCTTTCAATACCATCTTACCACAGAGCTATCCACAACTTCGCAGAG 780
 QY 905 TGTTTCAGAAATACCAAGTACAGTGAAGCATGACAAAGTTCTGACCTGGAGT 964
 Db 781 TGTTTCAGAAATACCAAGTACAGTGAAGCATGACAAAGTTCTGACCTGGAGT 840
 QY 965 TGAAGGTGACATTCAGAAATGCTCTCAGCTACAGCTATCGTGAAGTGCACCAAGCAAAACAG 1024
 Db 841 TGAAGGTGACATTCAGAAATGCTCTCAGCTACAGCTATCGTGAAGTGCACCAAGCAAAACAG 900
 QY 1025 CTCTTCGAGAGAGCTTCATCAGCCATGAAGGTTTGGAACTTCGCCATAGGCAT 1084
 Db 901 CTCTTCGAGAGAGCTTCATCAGCCATGAAGGTTTGGAACTTCGCCATAGGCAT 960
 QY 1085 TGATCAGGATTATGTTTCCCGTTCTGAAATTTGACATGAATGATCAAAAGCATTCATC 1144
 Db 961 TGATCAGGATTATGTTTCCCGTTCTGAAATTTGACATGAATGATCAAAAGCATTCATC 1020
 QY 1145 AGAAGATGATGTTATCTCCCTTTGCCAAGCCATCTCGGATGAACCAAGAGAGATG 1204
 Db 1021 AGAAGATGATGTTATCTCCCTTTGCCAAGCCATCTCGGATGAACCAAGAGAGATG 1080
 QY 1205 AGAATCTCTGCTGCTCTTTGTGAGAGAACTAAACATTCCTTGTGTTCTCAAGCTA 1264
 Db 1081 AGAATCTCTGCTGCTCTTTGTGAGAGAACTAAACATTCCTTGTGTTCTCAAGCTA 1140
 QY 1265 TGATCAGAGAGCTTTAATATATATTTTCACTTATAAGCTTAATAGGAAGTTCTTC 1324
 Db 1141 TGATCAGAGAGCTTTAATATATATTTTCACTTATAAGCTTAATAGGAAGTTCTTC 1200
 QY 1325 AACAGATTACAGTACGCTACATCTGAAATATAGCTTTTAAATCATTTTAT 1384
 Db 1201 AACAGATTACAGTACGCTACATCTGAAATATAGCTTTTAAATCATTTTAT 1260
 QY 1385 AATTAACCTCTCTATAATAGAGTAAGTCCATTTTAAAAATGTTTCCCAACACATA 1444
 Db 1261 AATTAACCTCTCTATAATAGAGTAAGTCCATTTTAAAAATGTTTCCCAACACATA 1320
 QY 1445 AAACCCCTATACAGTTGTTCTAGTAAACATATACATGAGAAAGTCTATGCTGAGAA 1504

Db 1321 AAACCCCTATACAGTTGTTCTAGTAAACAATACATGAGAAAGATGTTCTATGTAGCTGAAA 1380
 QY 1505 TAAATGNCCTC 1516
 Db 1381 TAAATGACGTC 1392
 RESULT 4
 AX281593 1399 bp DNA linear PAT 02-NOV-2001
 LOCUS Sequence 2 from Patent WO0177389.
 DEFINITION
 AX281593
 ACCESSION
 AX281593.1 GI:16608844
 VERSION
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 AUTHORS
 1 Shifman,D., Somogyi,R., Lawn,R., Seilhamer,J.J., Porter,G.J.,
 Mikita,T. and Tai,J.
 TITILE Genes expressed in foam cell differentiation
 JOURNAL Patent: WO 0177389-A 2 18-OCT-2001;
 Incyte Genomics, Inc. (US)
 FEATURES
 Location/Qualifiers
 source
 1. 1399
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 /note="Incyte ID No: g34387"
 ORIGIN
 Query Match 91.5%; Score 1387.8; DB 6; Length 1399;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1389; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 125 AGTGTGAATCTTCAGAGAAGATTTCTTCTTAGTCTTTTTCAGAAAGTAGAGATAAG 184
 Db 1 AGTGTGAATCTTCAGAGAAGATTTCTTCTTAGTCTTTTTCAGAAAGTAGAGATAAG 60
 QY 185 ACATTTTCAAAAATGGCAATGGTATCAGAAATTCCTCAAGCAGGCTGGTTATTGAA 244
 Db 61 ACATTTTCAAAAATGGCAATGGTATCAGAAATTCCTCAAGCAGGCTGGTTATTGAA 120
 QY 245 ATGAGAGCAGAAATGTTCAAACTGTGAAGTCAATCAAAAGGTGGTCCCGATCAGCGG 304
 Db 121 ATGAGAGCAGAAATGTTCAAACTGTGAAGTCAATCAAAAGGTGGTCCCGATCAGCGG 180
 QY 305 TGAGCCCTTATCTTACCTTCAATCCATCTCGATGTCGCTGCTTGCATAGGCCATAA 364
 Db 181 TGAGCCCTTATCTTACCTTCAATCCATCTCGATGTCGCTGCTTGCATAGGCCATAA 240
 QY 365 TGGTTAAAGGTGGATGAAGCAACCATTCATTAAGCAAGGAAAGCCCTGGATGAACAC 424
 Db 241 TGGTTAAAGGTGGATGAAGCAACCATTCATTAAGCAAGGAAAGCCCTGGATGAACAC 300
 QY 425 AGCGTCAACAGATCAAGCAGCATATCTCCAGAAACAGGAAAGCCCTGGATGAACAC 484
 Db 301 AGCGTCAACAGATCAAGCAGCATATCTCCAGAAACAGGAAAGCCCTGGATGAACAC 360
 QY 485 TGAAGAAAGCCCTTACAGGTCACTTGAGAGGTTGTTTAGCTGTCTGCTAAACCTCCAG 544
 Db 361 TTAAGAAAGCCCTTACAGGTCACTTGAGAGGTTGTTTAGCTGTCTGCTAAACCTCCAG 420
 QY 545 CGCAATTTGATGCTGATGAATCTGCTGCTCCATGAAGGCTTGGAACTGATGAAGATA 604
 Db 421 CGCAATTTGATGCTGATGAATCTGCTGCTCCATGAAGGCTTGGAACTGATGAAGATA 480
 QY 605 CTCTAATTTGAGATTGTCGATCAAGAACTAAACAAGAAATCAGAGCATTAACAGGCTCT 664
 Db 481 CTCTAATTTGAGATTGTCGATCAAGAACTAAACAAGAAATCAGAGCATTAACAGGCTCT 540
 QY 665 ACAGAGAGGAACTGAAGAGATCTGCGCAAGACATTAACCTCAGACATCTGGAGATT 724

Db	541	ACAGAGAGAACTGAAGAGAGATCTGGCCAAAGACATAAACCCTCAGACACATCTGGAGATT	600
Qy	725	TTCCGGAACGCTTTGCTTCTCTGTTAAGGGTGACCGGATCTGAGGACTTTGGTGTGAATG	784
Db	601	TTCCGGAACGCTTTGCTTCTCTGTTAAGGGTGACCGGATCTGAGGACTTTGGTGTGAATG	660
Qy	785	AAGACTTGCGTGAATTCAGATGCCAGGGCTTGTATGAAGCGAGGAGAAAGGAGGGA	844
Db	661	AAGACTTGCGTGAATTCAGATGCCAGGGCTTGTATGAAGCGAGGAGAAAGGAGGGA	720
Qy	845	CAGAGCTAAACGCTTTCAATACCATCTTACCACAGAGAGCTATCCACAACCTTCGCAGAG	904
Db	721	CAGAGCTAAACGCTTTCAATACCATCTTACCACAGAGAGCTATCCACAACCTTCGCAGAG	780
Qy	905	TGTTTCAGAAATACACCAAGTACAGTAAGCATGACATGAACAAAGTTCTGCACTTGGAGT	964
Db	781	TGTTTCAGAAATACACCAAGTACAGTAAGCATGACATGAACAAAGTTCTGCACTTGGAGT	840
Qy	965	TGAAGGTTGACATTCAGAAATGCTTCAGAGCTATCTGTAAGTGCGCCACAGCAACCCAG	1024
Db	841	TGAAGGTTGACATTCAGAAATGCTTCAGAGCTATCTGTAAGTGCGCCACAGCAACCCAG	900
Qy	1025	CTTTCTTTGCAGAGAGCTTCTCAAGCCATGAAAGGTTTGGAACTCGCCATAGGCAT	1084
Db	901	CTTTCTTTGCAGAGAGCTTCTCAAGCCATGAAAGGTTTGGAACTCGCCATAGGCAT	960
Qy	1085	TGATCAGGTTATGTTTCCCGTCTGAAATGACATGATATCAAGCATTTCTATC	1144
Db	961	TGATCAGGTTATGTTTCCCGTCTGAAATGACATGATATCAAGCATTTCTATC	1020
Qy	1145	AGAAGATGATGGTATCTCCCTTCCCAAGCCATCTCGATGAACACCAAGGAGATG	1204
Db	1021	AGAAGATGATGGTATCTCCCTTCCCAAGCCATCTCGATGAACACCAAGGAGATG	1080
Qy	1205	AGAAATCTGGTGGCTTTTGGAGGAACTAAACATTCCTTGATGGTCTCAAGCTA	1264
Db	1081	AGAAATCTGGTGGCTTTTGGAGGAACTAAACATTCCTTGATGGTCTCAAGCTA	1140
Qy	1265	TGATCAGAGACTTTAAATATATATTTTCATCTATAAGCTTAAATAGGAAAGTTTCTC	1324
Db	1141	TGATCAGAGACTTTAAATATATATTTTCATCTATAAGCTTAAATAGGAAAGTTTCTC	1200
Qy	1325	ACAGGATACAGTGTAGTACTACCTGCTGAAATAATAGCCCTTAAATCATTTTAT	1384
Db	1201	ACAGGATACAGTGTAGTACTACCTGCTGAAATAATAGCCCTTAAATCATTTTAT	1260
Qy	1385	ATTATACTCTGTATAATAGATAGTCCATTTTAAATAATGTTTCCCAACCCATA	1444
Db	1261	ATTATACTCTGTATAATAGATAGTCCATTTTAAATAATGTTTCCCAACCCATA	1320
Qy	1445	AAACCTATACAGTTGTTCTAGTAAACAATACATGAGAAAGATGCTATGTAGCTGAAA	1504
Db	1321	AAACCTATACAGTTGTTCTAGTAAACAATACATGAGAAAGATGCTATGTAGCTGAAA	1380
Qy	1505	TAAATGNGCTC	1516
Db	1381	TAAATGAGCTC	1392

RESULT 5
AX334637
LOCUS AX334637 1399 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 5146 from Patent WO0194629.
ACCESSION AX334637
VERSION AX334637.1 GI:18125356
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Young, P.B., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,

TITLE Horrigan, S., Soppet, D.R. and Weaver, Z.
Cancer gene determination and therapeutic screening using signature
Gene sets
JOURNAL Patent: WO 0194629-A 5146 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES Location/Qualifiers
source 1..1399
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN	Query Match	91.5%;	Score 1387.8;	DB 6;	Length 1399;
	Best Local Similarity	99.8%;	Pred. No. 0;		
	Matches 1389;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;
Qy	125	AGTGTGAATCTTTCAGAGAGAAATTTCTCTTTAGTTCTTTTGAAGAGGTAGAGATAAAG	184		
Db	1	AGTGTGAATCTTTCAGAGAGAAATTTCTCTTTAGTTCTTTTGAAGAGGTAGAGATAAAG	60		
Qy	185	ACATTTTTCAAAAATGGCAATGGTATCAGAAATTCCTCAAGCAGGCTGTTTATTGAAA	244		
Db	61	ACATTTTTCAAAAATGGCAATGGTATCAGAAATTCCTCAAGCAGGCTGTTTATTGAAA	120		
Qy	245	ATGAAGAGCAGGAATATGTTCAAACCTGTGAAGTCATCCAAAAGTGTCCCGGATCAGCG	304		
Db	121	ATGAAGAGCAGGAATATGTTCAAACCTGTGAAGTCATCCAAAAGTGTCCCGGATCAGCG	180		
Qy	305	TGAGCCCTTATCTTACCTTCAATCCATCTCCGATGCTCGCTTGCATTAAGCCATAA	364		
Db	181	TGAGCCCTTATCTTACCTTCAATCCATCTCCGATGCTCGCTTGCATTAAGCCATAA	240		
Qy	365	TGTTAAAGTGTGGATGAAGCAACCATCATTCACATTTTAACTAAAGCGAAACAATGCAC	424		
Db	241	TGTTAAAGTGTGGATGAAGCAACCATCATTCACATTTTAACTAAAGCGAAACAATGCAC	300		
Qy	425	AGCGTCAACAGATCAAGCAGCATATCTCCAGAAACAGAGAGCCCTGGATGAACAC	484		
Db	301	AGCGTCAACAGATCAAGCAGCATATCTCCAGAAACAGAGAGCCCTGGATGAACAC	360		
Qy	485	TGAAGAAAGCCCTTACAGGTCACCTTGAAGAGGTTGTTTGTAGCTCTGCTAAAACTCCAG	544		
Db	361	TGAAGAAAGCCCTTACAGGTCACCTTGAAGAGGTTGTTTGTAGCTCTGCTAAAACTCCAG	420		
Qy	545	CGMAATTTGATGTGATGAACCTTCTGCTGCCATGAAGGCCCTTGGAACTGTAGAGATA	604		
Db	421	CGMAATTTGATGTGATGAACCTTCTGCTGCCATGAAGGCCCTTGGAACTGTAGAGATA	480		
Qy	605	CTCTAATTGAGATTTTGGCATCAAGAACTAAACAAAGAAATCAGAGACATTTAACAGGGTCT	664		
Db	481	CTCTAATTGAGATTTTGGCATCAAGAACTAAACAAAGAAATCAGAGACATTTAACAGGGTCT	540		
Qy	665	ACAGAGGAACTGAAGAGAGATCTGGCCAAAGACATAACCTCAGACACATCTCGAGATT	724		
Db	541	ACAGAGGAACTGAAGAGAGATCTGGCCAAAGACATAACCTCAGACACATCTCGAGATT	600		
Qy	725	TTCCGGAACGCTTTGCTTCTCTGTTAAGGGTGACCGATCTGAGGACTTTGGTGTGAATG	784		
Db	601	TTCCGGAACGCTTTGCTTCTCTGTTAAGGGTGACCGATCTGAGGACTTTGGTGTGAATG	660		
Qy	785	AAGACTTTGGCTGATTCAGATGCCAGGGCTTGTATGAAGCAGGAGAAAGGAGAAAGGGA	844		
Db	661	AAGACTTTGGCTGATTCAGATGCCAGGGCTTGTATGAAGCAGGAGAAAGGAGAAAGGGA	720		
Qy	845	CAGAGCTAAACGCTTTCAATACCATCTTACCACAGAGAGCTATCCACAACCTTCGCAGAG	904		
Db	721	CAGAGCTAAACGCTTTCAATACCATCTTACCACAGAGAGCTATCCACAACCTTCGCAGAG	780		
Qy	905	TGTTTCAGAAATACACCAAGTACAGTAAGCATGACATGAACAAAGTTCTGCACTTGGAGT	964		
Db	781	TGTTTCAGAAATACACCAAGTACAGTAAGCATGACATGAACAAAGTTCTGCACTTGGAGT	840		
Qy	965	TGAAGGTTGACATTCAGAAATGCTTCAGAGCTATCTGTAAGTGCGCCACAGCAACCCAG	1024		

```
Db      841  TGAAGGTGACATTGAGAAATGCTCACAGCTATCTGTAAGTGGCGCAACAAGCAACAG 900
Qy      1025  CTTCTTTGAGAGAGCTTCATCAAGCCATGAAGGTGTTGGAAGTCCGCTCCATAGGAT 1084
Db      901  CTTCTTTGAGAGAGCTTCATCAAGCCATGAAGGTGTTGGAAGTCCGCTCCATAGGAT 960
Qy      1085  TGATCAGGATTATGGTTTCCGTTCTGAAATGACATGAATGATATCAAAAGCATCTATC 1144
Db      961  TGATCAGGATTATGGTTTCCGTTCTGAAATGACATGAATGATATCAAAAGCATCTATC 1020
Qy      1145  AGAAGATGATGTAATCTCCCTTGGCAAGCCATCTGGATGAACCAAGAGAGATG 1204
Db      1021  AGAAGATGATGTAATCTCCCTTGGCAAGCCATCTGGATGAACCAAGAGATG 1080
Qy      1205  AGAAATCTCGTGGTCTCTTGGAGGAACTAAACATTCCTCGATGATCTCAAGCTA 1264
Db      1081  AGAAATCTCGTGGTCTCTTGGAGGAACTAAACATTCCTCGATGATCTCAAGCTA 1140
Qy      1265  TGATCAGAGACTTAATATATATTTTCACTATAAGCTTAAGTAAAGATTTCTTC 1324
Db      1141  TGATCAGAGACTTAATATATATTTTCACTATAAGCTTAAGTAAAGATTTCTTC 1200
Qy      1325  AACAGGATTACAGGTAGCTACCTACATGCTGAAATATAGCTTTAAATCATTTTAT 1384
Db      1201  AACAGGATTACAGGTAGCTACCTACATGCTGAAATATAGCTTTAAATCATTTTAT 1260
Qy      1385  ATTATACTCTGTATAATAGAGATAAGTCCATTTTAAATGTTTCCCAACCATTA 1444
Db      1261  ATTATACTCTGTATAATAGAGATAAGTCCATTTTAAATGTTTCCCAACCATTA 1320
Qy      1445  AAACCTTATCAAGTGTCTTAGTAAACATACATGAGAAAGATGTCTATGAGCTGAAA 1504
Db      1321  AAACCTTATCAAGTGTCTTAGTAAACATACATGAGAAAGATGTCTATGAGCTGAAA 1380
Qy      1505  TAAATGNCGTC 1516
Db      1381  TAAATGACGTC 1392

RESULT 6
AX779855 1399 bp DNA linear PAT 14-JUL-2003
LOCUS     Sequence 2012 from Patent WO03039443.
DEFINITION
ACCESSION AX779855
VERSION   AX779855.1 GI:32696849
KEYWORDS  Homo sapiens (human)
SOURCE    Homo sapiens
ORGANISM  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1. Haerlach, T.; Schoch, C.; Kern, W.; Kohlmann, A.; Schnittger, S.;
    Dugas, M.; Ellis, R.; Brors, B. and Mergenthaler, S.
    Novel genetic markers for leukemias
    Patent: WO 03039443-A 2012 15-MAY-2003;
    Deutsches Krebsforschungszentrum (DKFZ);
    Ludwig-Maximilians-Universitaet Muenchen (LMU);
    PD Dr. Dr. (DB); Schoch, Claudia (DB); Kern, Wolfgang (DB)
FEATURES
    source
        1..1399
        /organism="Homo sapiens"
        /mol_type="unassigned DNA"
        /db_xref="taxon:9606"
ORIGIN
    Query Match 91.5%; Score 1387.8; DB 6; Length 1399;
    Best Local Similarity 99.8%; Pred. No. 0;
    Matches 1399; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy      125  AGTGTGAATCTTCAGAGAGAAATTTCTTAGTTCTTTCCAGAGAGGTAGATTAAG 184
Db      1  AGTGTGAATCTTCAGAGAGAAATTTCTTAGTTCTTTCCAGAGAGGTAGATTAAG 60
```

```
Qy      185  ACCTTTTCAAAAATGGCAATGGTATCAGAAATCTCTCAAGCAGGCTGTTATTGAAA 244
Db      61  ACCTTTTCAAAAATGGCAATGGTATCAGAAATCTCTCAAGCAGGCTGTTATTGAAA 120
Qy      245  ATGAAGAGCAGGAATATGTTCAAACTGTGAAGTCAATCAAAAGGTGGTCCCGGATCAGCGG 304
Db      121  ATGAAGAGCAGGAATATGTTCAAACTGTGAAGTCAATCAAAAGGTGGTCCCGGATCAGCGG 180
Qy      305  TGAGCCCTATCTTACCTTCAATCCATCTCGGATGTCGCTGCTTCATAGGCCATTA 364
Db      181  TGAGCCCTATCTTACCTTCAATCCATCTCGGATGTCGCTGCTTCATAGGCCATTA 240
Qy      365  TGGTTAAAGGTGGATGAAGCAACCATCATTTGACATTTCTAACTAAAGCAAAACAATCCAC 424
Db      241  TGGTTAAAGGTGGATGAAGCAACCATCATTTGACATTTCTAACTAAAGCAAAACAATCCAC 300
Qy      425  AGCGTCAACAGATCAAAAGCAGCATATCTCCAGGAAACAGGAAGCCCTCGATGAACAC 484
Db      301  AGCGTCAACAGATCAAAAGCAGCATATCTCCAGGAAACAGGAAGCCCTCGATGAACAC 360
Qy      485  TGAAGAAAGCCCTTACAGGTCACCTTGAGGAGGTTGTTTGTAGCTCTCTAAATACTCCAG 544
Db      361  TTAAGAAAGCCCTTACAGGTCACCTTGAGGAGGTTGTTTGTAGCTCTCTAAATACTCCAG 420
Qy      545  CGCAATTTGATGCTGATGAACCTTCTGCTGCAATGAAGGCGCTTTGGAACATGATGAAGATA 604
Db      421  CGCAATTTGATGCTGATGAACCTTCTGCTGCAATGAAGGCGCTTTGGAACATGATGAAGATA 480
Qy      605  CTCATATTTGATTTGGCATCAAGAACTAAAGAAAGAAATCAGAGACATTAACAGGGTCT 664
Db      481  CTCATATTTGATTTGGCATCAAGAACTAAAGAAAGAAATCAGAGACATTAACAGGGTCT 540
Qy      665  ACAGAGAGAACTGAAGAGAGATCTGGCCAAAGACATAACCTCAGACACATCTCGAGATT 724
Db      541  ACAGAGAGAACTGAAGAGAGATCTGGCCAAAGACATAACCTCAGACACATCTCGAGATT 600
Qy      725  TTCGAAACGCTTTGCTTCTTCTGCTTAAGGTCACCGATCTGAGGACTTTGGTGTGATG 784
Db      601  TTCGAAACGCTTTGCTTCTTCTGCTTAAGGTCACCGATCTGAGGACTTTGGTGTGATG 660
Qy      785  AAGACTTCGGCTGATTCAGATGCCAGGCGCTTGTATGAAGCAGGAGAAAGAGAGGGA 844
Db      661  AAGACTTCGGCTGATTCAGATGCCAGGCGCTTGTATGAAGCAGGAGAAAGAGAGGGA 720
Qy      845  CAGACGTAAACGCTGTTCAATACCATCTTACCACCAAGAGCTATCCACAACTTCGACAG 904
Db      721  CAGACGTAAACGCTGTTCAATACCATCTTACCACCAAGAGCTATCCACAACTTCGACAG 780
Qy      905  TGTTCAGAAATACACCAGTACAGTAAAGTACATCAACAAAGTTCCTGACCTCGAGT 964
Db      781  TGTTCAGAAATACACCAGTACAGTAAAGTACATCAACAAAGTTCCTGACCTCGAGT 840
Qy      965  TGAAGGTGACATTCAGAAATGCTTCACAGCTATCTGTAAGTGGCCACACAGCAACACAG 1024
Db      841  TGAAGGTGACATTCAGAAATGCTTCACAGCTATCTGTAAGTGGCCACACAGCAACACAG 900
Qy      1025  CTTCTTTGAGAGAGCTTCATCAAGCCATGAAGGTGTTGGAAGTCCGCTCCATAGGAT 1084
Db      901  CTTCTTTGAGAGAGCTTCATCAAGCCATGAAGGTGTTGGAAGTCCGCTCCATAGGAT 960
Qy      1085  TGATCAGGATTATGGTTTCCGTTCTGAAATGACATGAATGATATCAAGCATTTCTATC 1144
Db      961  TGATCAGGATTATGGTTTCCGTTCTGAAATGACATGAATGATATCAAGCATTTCTATC 1020
Qy      1145  AGAAGATGATGTAATCTCCCTTGGCAAGCCATCTGGATGAACCAAGAGAGATG 1204
Db      1021  AGAAGATGATGTAATCTCCCTTGGCAAGCCATCTGGATGAACCAAGAGATG 1080
Qy      1205  AGAAATCTCGTGGTCTCTTGGAGGAACTAAACATTCCTCGATGATCTCAAGCTA 1264
Db      1081  AGAAATCTCGTGGTCTCTTGGAGGAACTAAACATTCCTCGATGATCTCAAGCTA 1140
```



```
QY 1265 TGATCAGAGACTTTAAATATATATATTTTCATCTCTATTAAGCTTTAAATAGAGAGTTTCTTC 1324
Db 1141 TGATCAGAGACTTTAAATATATATATTTTCATCTCTATTAAGCTTTAAATAGAGAGTTTCTTC 1200
QY 1325 AACAGGATTACAGTGTAGCTTACCTACATGCTGAAAAATATAGCCCTTTAAATCATTTTAT 1384
Db 1201 AACAGGATTACAGTGTAGCTTACCTACATGCTGAAAAATATAGCCCTTTAAATCATTTTAT 1260
QY 1385 ATTATTAACCTGTATTAATAGAGATAGTCCATTTTAAATATGTTTTCCCAACCCATA 1444
Db 1261 ATTATTAACCTGTATTAATAGAGATAGTCCATTTTAAATATGTTTTCCCAACCCATA 1320
QY 1445 AAACCCCTATACAAAGTTGTTCTAGTAAACAATACATGAGAAAGATGTCATGTAGCTGAAAA 1504
Db 1321 AAACCCCTATACAAAGTTGTTCTAGTAAACAATACATGAGAAAGATGTCATGTAGCTGAAAA 1380
QY 1505 TAAATGNCGTC 1516
Db 1381 TAAATGACGTC 1392

RESULT 8
AX004167/c
LOCUS 1398 bp DNA linear PAT 24-AUG-2000
DEFINITION Sequence 4 from Patent WO9921980.
ACCESSION AX004167
VERSION AX004167.1 GI:9927716
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Georges, E. and Wang, Y.
TITLE P-40/annexin i and related proteins and their role in multilidrug
resistance
JOURNAL Patent: WO 9921980-A 4 06-MAY-1999;
GEORGES ELIAS (CA); UNIV MCGILL (CA)
FEATURES
source
1. 1398
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 90.8%; Score 1375.8; DB 6; Length 1398;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1388; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
QY 125 AGTGTGAAATCTTCAGAGAGAAATTTCTCTTAGTTCTTTTCCAAAGAGGTAGAGATAAAG 184
Db 1398 AGTGTGAAATCTTCAGAGAGAAATTTCTCTTAGTTCTTTTCCAAAGAGGTAGAGATAAAG 1339
QY 185 ACATTTTCAAAATGGCAATGTATCAGAAATTCCTCAAGCAGGCGCTGTTTATTGAAA 244
Db 1338 -CACTTTTCAAAATGGCAATGTATCAGAAATTCCTCAAGCAGGCGCTGTTTATTGAAA 1280
QY 245 ATGAGACGAGATATGTTCCAACTGTGAAGTTCATCCAAAGGTGGTCCCGGATCAGCGG 304
Db 1279 ATGAGACGAGATATGTTCCAACTGTGAAGTTCATCCAAAGGTGGTCCCGGATCAGCGG 1220
QY 305 TGAGCCCTATCTCACTTCAATCTCGGATGTCGTCGCTTGCATTAAGGCCATAA 364
Db 1219 TGAGCCCTATCTCACTTCAATCTCGGATGTCGTCGCTTGCATTAAGGCCATAA 1160
QY 365 TGGTTAAAGGTGGATGACGACCACTCATTTGACATTTCTAATCAAGCAACCAATGCAC 424
Db 1159 TGGTTAAAGGTGGATGACGACCACTCATTTGACATTTCTAATCAAGCAACCAATGCAC 1100
QY 425 AGCGTCAACAGATCAAGACGACATATCTCCAGGAAACAGGAAGCCCTGGATGAACAC 484
Db 1099 AGCGTCAACAGATCAAGACGACATATCTCCAGGAAACAGGAAGCCCTGGATGAACAC 1040
```

```
QY 485 TGAAGAAAGCCCTTACAGGTCACTTGGAGAGGTGTGTTTAGCTCTGCTAAAAACCCAG 544
Db 1039 TTAAGAAAGCCCTTACAGGTCACTTGGAGAGGTGTGTTTAGCTCTGCTAAAAACCCAG 980
QY 545 CGCAATTTGATGTGATGAACTTCCTGCTGCTCCATGAAGGSCCTTGGAACTGATGAAGTA 604
Db 979 CGCAATTTGATGTGATGAACTTCCTGCTGCTCCATGAAGGSCCTTGGAACTGATGAAGTA 920
QY 605 CTCTAAATTTGAGATTTTGGCATCAAGAACTAAACAAAGAAATCAGAGACATTTAAACAGGCTCT 664
Db 919 CTCTAAATTTGAGATTTTGGCATCAAGAACTAAACAAAGAAATCAGAGACATTTAAACAGGCTCT 860
QY 665 ACAGAGAGAACTGAAGAGAGATCTGSCCAAGACATTAACCTCAGACACATCTCGAGATT 724
Db 859 ACAGAGAGAACTGAAGAGAGATCTGSCCAAGACATTAACCTCAGACACATCTCGAGATT 800
QY 725 TTCGGAACGCTTTGCTTTCTCTTAAGGCTGACCGATCTGAGGACTTTTGGTGTGAATG 784
Db 799 TTCGGAACGCTTTGCTTTCTCTTAAGGCTGACCGATCTGAGGACTTTTGGTGTGAATG 740
QY 785 AAGACTTTGGCTGATTCAGATGCCAGGSCCTTGTATGAAGCAGAGAGAAAGGAGAAAGGGA 844
Db 739 AAGACTTTGGCTGATTCAGATGCCAGGSCCTTGTATGAAGCAGAGAGAAAGGAGAAAGGGA 680
QY 845 CAGACGTAAAGCTGTTCAATACCATCTTACCACAGAGCTATCCACAACTTCGCAGAG 904
Db 679 CAGACGTAAAGCTGTTCAATACCATCTTACCACAGAGCTATCCACAACTTCGCAGAG 620
QY 905 TGTTTTCAGAAATACACCAAGTACAGTAAGCATGACATGAACAAAGTTCTGGACCTGAGT 964
Db 619 TGTTTTCAGAAATACACCAAGTACAGTAAGCATGACATGAACAAAGTTCTGGACCTGAGT 560
QY 965 TGAAGGTGACATTCAGAAATGCTCCTCAGCTATCGTGAAGTGGCCACAGCAAAACCCAG 1024
Db 559 TGAAGGTGACATTCAGAAATGCTCCTCAGCTATCGTGAAGTGGCCACAGCAAAACCCAG 500
QY 1025 CTTTCTTTGACAGAGAGCTTCATCAAGCCATGAAAGTGTGTGAATCTGCGCATAGGCAT 1084
Db 499 CTTTCTTTGACAGAGAGCTTCATCAAGCCATGAAAGTGTGTGAATCTGCGCATAGGCAT 440
QY 1085 TGATCAGATTTATGTTTCCGTTCTGAAATGACATGAATGATATCAAAAGCATTTCTATC 1144
Db 439 TGATCAGATTTATGTTTCCGTTCTGAAATGACATGAATGATATCAAAAGCATTTCTATC 380
QY 1145 AGAAGATGTATGTTATCTCCCTTTGCCAAGCCATCTCGATGAAACCAAGAGAGATG 1204
Db 379 AGAAGATGTATGTTATCTCCCTTTGCCAAGCCATCTCGATGAAACCAAGAGAGATG 320
QY 1205 AGAAATCCTGCTGCTCTTTGTGAGGAAACTAAACATTCCTTGTATGTTCTCAAGCTA 1264
Db 319 AGAAATCCTGCTGCTCTTTGTGAGGAAACTAAACATTCCTTGTATGTTCTCAAGCTA 260
QY 1265 TGATCAGAGAGCTTTTAAATATATATATTTTTCATCTTAAGCTTAAATAGGAAAGTTTCTTC 1324
Db 259 TGATCAGAGAGCTTTTAAATATATATATTTTTCATCTTAAGCTTAAATAGGAAAGTTTCTTC 200
QY 1325 AACAGGATTACAGTGTAGCTTACCTACATGCTGAAATAATATAGCCCTTTAAATCATTTTAT 1384
Db 199 AACAGGATTACAGTGTAGCTTACCTACATGCTGAAATAATATAGCCCTTTAAATCATTTTAT 140
QY 1385 ATTATAAATCTCTGATAATAGAGATAAGTCCATTTTAAATAATGTTTTTCCCAACCCATA 1444
Db 139 ATTATAAATCTCTGATAATAGAGATAAGTCCATTTTAAATAATGTTTTTCCCAACCCATA 80
QY 1445 AAACCCCTATACAAAGTTGTTCTAGTAAACAATACATGAGAAAGATGTCATGTAGCTGAAAA 1504
Db 79 AAACCCCTATACAAAGTTGTTCTAGTAAACAATACATGAGAAAGATGTCATGTAGCTGAAAA 20
QY 1505 TAAATGNCGTC 1516
Db 19 TAAATGACGTC 8
```

RESULT 9	BC001275	1408 bp	mRNA	linear	PRI 04-OCT-2003
LOCUS	Homo sapiens annexin A1, mRNA (cDNA clone MGC:5095 IMAGE:3459615), complete cds.				
DEFINITION	BC001275.1 GI:12654862				
ACCESSION	MGC				
VERSION	Homo sapiens (human)				
KEYWORDS	Homo sapiens				
SOURCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
ORGANISM	1 (bases 1 to 1408)				
REFERENCE	Straussberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Kleitman, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schaller, G.D., Altschul, S.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, P., Hopkins, R.F., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Scheetz, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, R.J., Malek, J.A., Gunaratne, P.H., Richards, S., Wille, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Vollick, D.K., Mazny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shvchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickinson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalios, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.				
AUTHORS	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences				
TITLE	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)				
JOURNAL	2238257				
MEDLINE	12477932				
PUBMED	2 (bases 1 to 1408)				
REFERENCE	Straussberg, R.				
AUTHORS	Submitted (11-DEC-2000) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA				
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov				
COMMENT	Contact: MGC help desk Email: cgapbs-r@mail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305 Web site: http://www.shgc.stanford.edu Contact: (Dickson, Mark) mcdpaxil@stanford.edu Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.				
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAX Plate: 4 Row: 9 Column: 18 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4502100. Location/Qualifiers 1. .1408 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="MGC:5095 IMAGE:3459615" /issue_type="Cervix, carcinoma" /clone_lib="NIH MGC_12" /lab_host="DH10B" /note="Vector: pCMV-SPORT6" 1. .1408 /gene="ANXA1"				
source					
gene					

QY	748	GCTAAGGGTGACCGATCTGAGGACTTTTGGTGTGAATTAAAGACTTCGGCTGATTTCAGATGCC	807
Dd	601	GCTAAGGGTGACCGATCTGAGGACTTTTGGTGTGAATTAAAGACTTCGGCTGATTTCAGATGCC	660
QY	808	AGGCGCTTGTATGAAGCAGAGAAAAGGAGAAGGGGACAGACGTAAAAAGTGTTCNAATACC	867
Dd	661	AGGCGCTTGTATGAAGCAGAGAAAAGGAGAAGGGGACAGACGTAAAAAGTGTTCNAATACC	720
QY	868	ATCTCTTACCACCAAGAACTATCCACAACATTCGCAGAGTGTTCAGAAAAATACACCAAGTAC	927
Dd	721	ATCTCTTACCACCAAGAACTATCCACAACATTCGCAGAGTGTTCAGAAAAATACACCAAGTAC	780
QY	928	AGTAAGCATGACATGAACAAGTTCGGACCTGGAGTTCGAAAGGTGACATTCAGAAAAATGC	987
Dd	781	AGTAAGCATGACATGAACAAGTTCGGACCTGGAGTTCGAAAGGTGACATTCAGAAAAATGC	840
QY	988	CTCACAGCTATCGTGAAGTTCGCCACAAGCAAACCAAGCTTTCTTTGCAGAGAAAGCTTCAT	1047
Dd	841	CTCACAGCTATCGTGAAGTTCGCCACAAGCAAACCAAGCTTTCTTTGCAGAGAAAGCTTCAT	900
QY	1048	CHAAGCCATGAAGGTGTGAACTCCGCCATAAGGCATTCATCAGGATTCATGTTTCCCGT	1107
Dd	901	CHAAGCCATGAAGGTGTGAACTCCGCCATAAGGCATTCATCAGGATTCATGTTTCCCGT	960
QY	1108	TCTGAAATTCGATGAATGATATCAAGCAATTCATCAGAAGATGATGGTATCTCCCCT	1167
Dd	961	TCTGAAATTCGATGAATGATATCAAGCAATTCATCAGAAGATGATGGTATCTCCCCT	1020
QY	1168	TGCCAACCCATCTCGGATGAAACCCAAAGGAGAGTATGAGAAAAATCCTGGTGCTCTTTGT	1227
Dd	1021	TGCCAACCCATCTCGGATGAAACCCAAAGGAGATATGAGAAAAATCCTGGTGCTCTTTGT	1088
QY	1228	GGAGAAACTAAACATTCCTTGATGTCTCAAGCTATGATCAGAAGACTTTAAATATAT	1288
Dd	1081	GGAGAAACTAAACATTCCTTGATGTCTCAAGCTATGATCAGAAGACTTTAAATATAT	1144
QY	1288	ATTTCATCTCTAATAGCTTAAATAGGAAAGTTTCTTCAACAGGATTCAGTGTAGCTACC	1347
Dd	1141	ATTTCATCTCTAATAGCTTAAATAGGAAAGTTTCTTCAACAGGATTCAGTGTAGCTACC	1201
QY	1348	TACATCTGAAAAATATAGCTTTAAATCAATTTTTATATATATTAATCTGTATATAGAGA	1407
Dd	1201	TACATCTGAAAAATATAGCTTTAAATCAATTTTTATATATTAATCTGTATATATAGAGA	1266
QY	1408	TAGTCCATTTTTTAAAAATGTTTTCCCCAAACCCATAAAACCCCTATACAAAGTTGTTCTAG	1466
Dd	1261	TAGTCCATTTTTTAAAAATGTTTTCCCCAAACCCATAAAACCCCTATACAAAGTTGTTCTAG	1321
QY	1468	TACAAATACATGAGAAAGATCTATAGCTGAAAAATAAAATGNCGTC	1516
Dd	1321	TACAAATACATGAGAAAGATCTATAGCTGAAAAATAAAATGNCGTC	1369
<hr/>			
RESULT_13			
BC035993		1385 bp	mRNA linear PRI 07-OCT-2000
LOCUS		Homo sapiens annexin A1, mRNA (cDNA clone MGC:32774 IMAGE:4662939)	
DEFINITION		complete cds.	
ACCESSION		BC035993	
VERSION		BC035993.1	GI:23958903
KEYWORDS		MGC.	
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
		(bases 1 to 1385)	
		Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,	
		Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,	
		Ahtshul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,	
		Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,	
		Diatchenko,I., Marusins,K., Farmer,A.A., Rubin,G.M., Hong,L.,	
		Caplanenko,M., Soares,M.B., Bonaldo,M.P., Casavant,T.L.,	

Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.W., Woreley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzyzanski, M.I., Skalska, U., Smalls, D.E., Schnerker, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22388257

1247732

2 (bases 1 to 1385)

Strausberg, R.

Direct Submission

Submitted (31-JUL-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgabs-remail.nih.gov

Tissue Procurement: CLONTECH

cDNA Library Preparation: CLONTECH Laboratories, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www-shgc.stanford.edu>

Contact: (Dickson, Mark) mcdepaxil.stanford.edu

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAL Plate: 41 Row: h Column: 9.

Location/Qualifiers

1. 1385

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:32774 IMAGE:4662939"

/tissue_type="Lung"

/clone_lib="NIH_MGC_77"

/lab_host="DH10B"

/notes="vector: pDNR-LIB"

1. 1385

/genes="ANKA1"

/notes="synonyms: ANK1, LPC1"

/db_xref="LocusID:301"

/db_xref="MIM:151690"

75_1115

/codon_start=1

/product="annexin 1"

/protein_id="AAH35993.1"

/db_xref="GI:23958904"

/db_xref="LocusID:301"

/translation="MANVSEFLKQAWFIENRQEVYQTVKSGKPGSAVSPYPTNPSSDVAALHKAIVKGVDEATIIDILTKRNAORQIQKAAVLETKPLDELTKKALGLKEVLVALLKTPAQFDADLRAMKGLTDEDTLIEILASRTNKEPLDINRVYREILKPLADITSDTSGDFRNALLSLAKGDSDFGVNEDLADSDALYEAGERRAGTGVNVENTILTRSPQLGRVRFQKTKYSGKDMNVLDLELNGDIEKCLTAIVFCATSKPAFFAKUHQAKWGKGVTRHKALIRIMVSRSEIDMNDIKAFYQKQYGISLCQAILDSTKGDYEKILVALCGGN"

213_407

/notes="annexin; Region: Annexin. This family of annexins also includes giardin that has been shown to function as an annexin"

/db_xref="CDD:pfam00191"

misc_feature 423..623

/notes="annexin; Region: Annexin. This family of annexins also includes giardin that has been shown to function as an annexin"

/db_xref="CDD:pfam00191"

587..875

misc_feature

/notes="annexin; Region: Annexin. This family of annexins also includes giardin that has been shown to function as an annexin"

/db_xref="CDD:pfam00191"

900..1100

misc_feature

/notes="annexin; Region: Annexin. This family of annexins also includes giardin that has been shown to function as an annexin"

/db_xref="CDD:pfam00191"

ORIGIN

Query Match 89.6%; Score 1357.8; DB 9; Length 1385;

Best Local Similarity 98.8%; Pred. No. 2.1e-313;

Matches 1368; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 125 AGTGTGAATCTTCAGAGAAGAAATTCCTTTAGTCTTTTCAAGAGAGGTAGAGATAAG 184

Db 1 AGTGTGAATCTTCAGAGAAGAAATTCCTTTAGTCTTTTCAAGAGAGGTAGAGATAAG 60

Qy 185 ACACTTTTCAAAAATGGCAATGGTATCAGAAATTCCTCAAGCAGGCCCTGGTTATTGAAA 244

Db 61 ACACTTTTCAAAAATGGCAATGGTATCAGAAATTCCTCAAGCAGGCCCTGGTTATTGAAA 120

Qy 245 ATCAAGAGCAGGAATATGTTCAACTGTGAAGTATCAAAAGGTGTCCTCCCGATCAGGG 304

Db 121 ATCAAGAGCAGGAATATGTTCAACTGTGAAGTATCAAAAGGTGTCCTCCCGATCAGGG 180

Qy 305 TGAGCCCTATCTCACTTCAATCCATCTCGATGTCGCTGCTTGCATTAAGGCCATAA 364

Db 181 TGAGCCCTATCTCACTTCAATCCATCTCGATGTCGCTGCTTGCATTAAGGCCATAA 240

Qy 365 TGCTTAAAGTGTGATGAGCAACCATCTTAATTAAGTAAAGCAACATTCGAC 424

Db 241 TGCTTAAAGTGTGATGAGCAACCATCTTAATTAAGTAAAGCAACATTCGAC 300

Qy 425 AGCGTCAACAGATCAAGCAGCATATCTCCAGGAAACAGGAAAGCCCTGGATGAACAC 484

Db 301 AGCGTCAACAGATCAAGCAGCATATCTCCAGGAAACAGGAAAGCCCTGGATGAACAC 360

Qy 485 TGAAGAAGCCCTTACAGTCACCTTGAGGAGGTGTTTGTAGCTCTGCTAAACTCCAG 544

Db 361 TGAAGAAGCCCTTACAGTCACCTTGAGGAGGTGTTTGTAGCTCTGCTAAACTCCAG 420

Qy 545 CGCAATTGATGCTGATGAATCTTGTGCTGCTGATGAAGGCTTGGAACTGATGAAGATA 604

Db 421 CGCAATTGATGCTGATGAATCTTGTGCTGCTGATGAAGGCTTGGAACTGATGAAGATA 480

Qy 605 CTCTAATTGAGATTTGGCATCAAGAACTTAACAGAAATCAGAGACATTAACAGGGTCT 664

Db 481 CTCTAATTGAGATTTGGCATCAAGAACTTAACAGAAATCAGAGACATTAACAGGGTCT 540

Qy 665 ACAGAGAGAACTGAAGAGAGATCTGGCCAAAGACATTAACCTCAGACATCTCGAGATT 724

Db 541 ACAGAGAGAACTGAAGAGAGATCTGGCCAAAGACATTAACCTCAGACATCTCGAGATT 600

Qy 725 TTCGGAACGCTTTGCTTCTTCTTGAAGGCTGACCGATCTGAGGACTTTGGTGTGAATG 784

Db 601 TTCGGAACGCTTTGCTTCTTCTTGAAGGCTGACCGATCTGAGGACTTTGGTGTGAATG 660

Qy 785 AGACTTGTGCTGATTCAGATGCGGGCTTGTATGAAGCAGGAGAAAGGAGAAAGGGA 844

Db 661 AGACTTGTGCTGATTCAGATGCGGGCTTGTATGAAGCAGGAGAAAGGAGAAAGGGA 720

Qy 845 CAGACTTAAAGCTGTTCATTAATCCATCTTACACAGAGCTTATCCCACTTCGACAG 904

Db 721 CAGACTTAAAGCTGTTCATTAATCCATCTTACACAGAGCTTATCCCACTTCGACAG 780

misc_feature


```

QY 963 GTTGAAGGTGACATTGAGAAATGCTTCACAGCATCTCGTGAAGTGGCCACAGCAACC 1022
Db 1647 GTTGAAGGTGACATTGAGAAATGCTTCACAGCATCTCGTGAAGTGGCCACAGCAACC 1706
QY 1023 AGCTTTCTTTGCGAGAGAGCTTCATCAAGCCATGAAGGTGTGAGTCCGCAATAGGC 1082
Db 1707 AGCTTTCTTTGCGAGAGAGCTTCATCAAGCCATGAAGGTGTGAGTCCGCAATAGGC 1766
QY 1083 ATTGATCAGGATTATGTTTCCCGTTCTGAAATGACATGAATGATCAAAAGCATTCTA 1142
Db 1767 ATTGATCAGGATTATGTTTCCCGTTCTGAAATGACATGAATGATCAAAAGCATTCTA 1826
QY 1143 TCAGAGATGTATGATATCTCCCTTTGCGAGAGAGCTTCGATGAACCAAGAGAGATA 1202
Db 1827 TCAGAGATGTATGATATCTCCCTTTGCGAGAGAGCTTCGATGAACCAAGAGAGATA 1886
QY 1203 TGAGAAATCTCGTGGCTCTTTGTGAGAGAACTAAACATTCCTTTGATGTCTCAAGC 1262
Db 1887 TGAGAAATCTCGTGGCTCTTTGTGAGAGAACTAAACATTCCTTTGATGTCTCAAGC 1946
QY 1263 TATGATCAGAGACTTAAATATATATTTTTCATCCTATAGCTTAAATAGGAAAGTTTCT 1322
Db 1947 TATGATCAGAGACTTAAATATATATTTTTCATCCTATAGCTTAAATAGGAAAGTTTCT 2006
QY 1323 TCACAGGATTACAGTGTAGCTACCTACATGCTGAAATAATAGCTTTTAAATCATTTTT 1382
Db 2007 TCACAGGATTACAGTGTAGCTACCTACATGCTGAAATAATAGCTTTTAAATCATTTTT 2066
QY 1383 ATATTAACTCTGTAATAGAGATAGTCCATTTTTTAAATAGTCTTCCCAAGCA 1442
Db 2067 ATATTAACTCTGTAATAGAGATAGTCCATTTTTTAAATAGTCTTCCCAAGCA 2126
QY 1443 TAAACCTTATACAGTGTCTTGTAGTAACATACATGAGAAAGATGTCTATGATAGTGA 1502
Db 2127 TAAACCTTATACAGTGTCTTGTAGTAACATACATGAGAAAGATGTCTATGATAGTGA 2186
QY 1503 AATAAATGACGTC 1516
Db 2187 AATAAATGACGTC 2200

RESULT 15
AF544227
LOCUS AF544227 1252 bp mRNA linear MAM 03-OCT-2002
DEFINITION Equus caballus lipocortin-1 mRNA, complete cds.
ACCESSION AF544227
VERSION AF544227.1 GI:23477705
KEYWORDS
SOURCE
ORGANISM Equus caballus (horse)
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
REFERENCE 1 (bases 1 to 1252)
Bryant, C.E., Allen, A. and Maskell, D.J.
Cloning of equine lipocortin-1 and its full cDNA sequence
JOURNAL Unpublished
AUTHORS Bryant, C.E., Allen, A. and Maskell, D.J.
TITLE Direct Submission
REFERENCE 2 (bases 1 to 1252)
Bryant, C.E., Allen, A. and Maskell, D.J.
Submitted (09-SEP-2002) Clinical Veterinary Medicine, University of
Cambridge, Madingley Road, Cambridge CB10 1B, UK
FEATURES
source
1..1252
Location/Qualifiers
/organism="Equus caballus"
/mol_type="mRNA"
/db_xref="taxon:9796"
168..1208
/notes="annexin 1"
/codon_start=1
/product="lipocortin-1"
/protein_id="AAN34819.1"
/db_xref="GI:23477706"
CDS

```

```

/translation="MSMVSFLKQAWFIENBEOEYIKAVKSGKGGPSAVSPYSPFNP
SSDVDAHLKAITVKGVDKATIIELITKNNARQOIKAAYLQEKPKDDEALKKALTG
HLEDVALAKITTPARFDGIDELRAAMKGLGTDEDTLIEILTSRTNKEIREINRVVREEL
KRDIAKDJTSDTSGDFQKALLSLAKGSDSEDFGNDLADSDAPALYVAGERRKGTGV
NFVNTILTRSYPHLRUPQMTYKYSKDMKVLDEMKGVNENFTAIIVKCAKSKPM
FPAELHNAWKAGTRDKILIRIMVSRSEVDNNDIKACYQKLYGISLCCQALIDETKGD
YEKILVALCGRD"
Query Match 64.4%; Score 976; DB 4; Length 1252;
Best Local Similarity 88.3%; Pred. No. 2e-222;
Matches 1072; Conservative 0; Mismatches 140; Indels 2; Gaps 1;
ORIGIN
QY 68 GCTAGTGTGGCTTCCTTTAAATCCTATATAATCAGAAAGCCCAAGCTCTCCTCCAGT 127
Db 37 GCTGGGTGGCTTCCTTTAAATCCTATATAATCAGAAAGCCCAAGCTCTCCTCCAGT 96
QY 128 GTGAATCTTTCAGAGAGAAATTTCTTTAGTTCCTTTGCAAGAAGGTAGAGATAAGACA 187
Db 97 GTGAACCCCTCAGAGAGAAACCTCTCTCAGTTCCTTTGGAAGACGGTGGAAAAATATA 156
QY 188 CTTTTCRAAATGGCAATGGTATCAGAAATTCCTCAAGCAGCCCTGGTTTATTGAAATG 247
Db 157 CTTTTCRAAATGGTCAATGGTATCAGAAATTCCTCAAGCAGCCCTGGTTTATTGAAATG 216
QY 248 AAGAGCAGGAATATGTTCAAACTGTGAAGTCTATCAAAAGGTGGTCCCGGATCAGCGTGA 307
Db 217 AAGAGCAGGAATATGTTCAAACTGTGAAGTCTATCAAAAGGTGGTCCCGGATCAGCGTGA 276
QY 308 GCCCTATCTACCTTCATTCATCTCTGGATGTCCTGGCTTGCATAGGCCATATGG 367
Db 277 GCCCTATCTACCTTCATTCATCTCTGGATGTCCTGGCTTGCATAGGCCATATGG 336
QY 368 TTAAGGTGTGGATGAAGCAACCATCATTTGAGATTTCTAACGAGGAAACAAATGCAACG 427
Db 337 TGAAGGTGTGGATGAAGCAACCATCATTTGAGATTTCTAACGAGGAAACAAATGCAACG 396
QY 428 GTCAACAGATCAAGCAGCATATCTCCAGAAACAGGAAAGCCCTGGATGAAACACTGA 487
Db 397 GTCAACAGATCAAGCAGCATATCTCCAGAAACAGGAAAGCCCTGGATGAAACACTGA 456
QY 488 AGAAGCCCTTACAGGTCACTTGAAGAGTGTGTTTGTAGCTCTGCTAAAAACTCCAGCGC 547
Db 457 AGAAGCCCTTACAGGTCACTTGAAGAGTGTGTTTGTAGCTCTGCTAAAAACTCCAGCGC 516
QY 548 AATTGATCTGATGAACTTCTGCTGCTGATGAGAGGCTTGGAACTGATGAGATATGC 607
Db 517 GGTTCATCTGATGAACTTCTGCTGCTGATGAGAGGCTTGGAACTGATGAGATATGC 576
QY 608 TAATTCAGATTTTGGCATCAAGAACTAAAGAAATCAGAGACATTAACAGGCTCTACA 667
Db 577 TGATTAATTTTGGCATCAAGAACTAAAGAAATTAAGGAAATTAACAGAGTCTATA 636
QY 668 GAGAGAACTGAAGAGAGATCTGGCCCAAGACATACCTCAGACACATCTCGAGATTTTC 727
Db 637 GAGAGAACTGAAGAGAGATCTGGCCCAAGACATACCTCAGACACATCTCGAGATTTTC 696
QY 728 GGAAAGCTTTGCTTTCTCTTCTAAAGGTGACCGATCTGAGGACTTTTGGTGTGAATGAG 787
Db 697 AGAAGCTTTGCTTTCTCTTCTAAAGGTGACCGATCTGAGGACTTTTGGTGTGAATGAG 756
QY 788 ACTTGGCTGATTCAGATCCCGGCTTGTATGAGCAGGAGAGAGAGAGAGAGAGAGAG 847
Db 757 ACTTGGCTGATTCAGATCCCGGCTTGTATGAGCAGGAGAGAGAGAGAGAGAGAGAGAG 816
QY 848 AGCTAAAGCTGTTCAATACCATCTTACCAAGAGCTATCCCACTTCGCGAGAGTGT 907
Db 817 ATGTGAACCTGTTTCAATACCATCTTACCAAGAGCTATCCCACTTCGCGAGAGTGT 876
QY 908 TTCAGAAATACCAAGTACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 967
Db 877 TTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 936

```

Qy	968	AAGTGCACATTGAGAAATGCCTCACAGCTATCGTGAAGTGCOCCTCAAGCAAAACCAAGCTT	1027
Db	937	AGGGGACGTCGAGAACTGCTTCACAGCTATTGTGAAGTGTGCGACAAGCAAAACCAATGT	996
Qy	1028	TCTTTCAGAGAGCTTCATCAAGCCATGAAAGGTGTTGGAACTGCCATAGGCATTGA	1087
Db	997	TCTTTCAGAGAGCTTCATCAAGCCATGAAAGGTGTTGGAACTGCCATAGGCATTGA	1056
Qy	1088	TCAGGATTATGGTTTCCCGTTCTTGAATTCACATGAATGATATCAAGCATTTCTATCAGA	1147
Db	1057	TCAGGATTATGGTTTCCCGTTCTTGAATTCACATGAATGATATCAAGCATTTCTATCAGA	1116
Qy	1148	AGATGATATGATATCTCCCTTTCGCAAGCCATCCTGGATGAAACCAAGGAGATATGAGA	1207
Db	1117	AGTGTATGATATCTCACTCTGCAAGCCATCCTGGATGAAACCAAGGAGATTTACGAAA	1176
Qy	1208	AAATCCTGGTGGCTCTTTGTGGAGGAACTTAAACATTTCCCTTTGATGGTCTCAAGC--TAT	1265
Db	1177	AGATCCTGGTGGCTCTCTGTGGAGAGACTTAAACATTTCCCTTTCAAGCATTTT	1236
Qy	1266	GATCAGAGAGCTTT	1279
Db	1237	CATCAGAGAGCTTT	1250

Search completed: March 11, 2004, 03:35:30
Job time : 3888.73 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

DM nucleic - nucleic search, using sw model
Run on: March 10, 2004, 08:01:57 ; Search time 368.965 seconds
(without alignments)
17454.974 Million cell updates/sec

Title: US-10-084-817-8
Perfect score: 1516
Sequence: 1 cttgttttggacatagct.....gctgaaataaaatgncgtc 1516

Scoring table: IDENTITY NUC
Gap 10.0, Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_29Jan04.*
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002s.*
7: Geneseqn2003as.*
8: Geneseqn2003bs.*
9: Geneseqn2003cs.*
10: Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1515	99.9	1516	6	ABs62753 Prostate
2	1515	99.9	1516	7	ABx77579 Different
3	1497.4	98.8	1517	4	ABa09124 Human lip
4	1495.6	98.7	2334	4	Aah72822 Human cer
5	1495.6	98.7	2334	4	Aah72822 Human cer
6	1389.4	91.5	1399	6	ABx09784 Human ova
7	1387.8	91.5	1399	2	AAX57357 Human p-4
8	1387.8	91.5	1399	6	ABl66809 Lung canc
9	1387.8	91.5	1399	6	AAs94747 Human DNA
10	1387.8	91.5	1399	6	ABx84511 Human cDN
11	1387.8	91.5	1399	6	ABv78074 Hypoxia-r
12	1379.4	91.0	1431	3	AAC77652 Human can
13	1367.4	90.2	1377	9	ADD18446 Human pro
14	1364.8	90.0	1376	1	AAN60555 Sequence
15	1364.8	90.0	1376	2	AAQ23216 Lambda IC
16	1358.4	89.6	1376	2	AAQ05809 Human lip
17	1309	86.3	1325	1	AAN60556 Sequence
18	1262.2	83.3	1432	6	ABz11387 Human pol
19	886.8	58.5	1402	7	ABK63678 Rat seque
20	886.8	58.5	1402	7	ABT41862 Toxicity
21	886.8	58.5	1402	9	ADB58147 Toxicity
22	880	58.0	1355	1	AAN82025 Fragment
23	879.4	58.0	1933	1	AAN90120 Human lip

24	772.2	50.9	1211	6	ABz11388
25	652.8	43.1	730	6	ABz11386
26	589.4	38.9	614	6	ABK09704 Human ova
27	566	37.3	567	3	AAC00928 Human sec
28	520.4	34.3	550	5	AaH82515 Human ova
29	519.4	34.3	537	4	AAs23953 Human ova
30	488.6	32.2	596	6	ABG59522 Human col
31	478.4	31.6	498	6	ABQ60100 Human col
32	459.4	30.3	461	6	ABV87296 Human col
33	457.8	30.2	461	6	ABV87296 Human col
34	444	29.3	683	6	ABQ54881 Human ova
35	433	28.6	434	8	ACH29835 Human tes
36	426.4	28.1	485	8	ACH20064 Human adu
37	424.4	28.0	438	6	AAs61555 Lung smal
38	387.6	25.6	505	8	ACH37092 Human end
39	362.8	23.9	542	2	Az117442 Human gen
40	349.4	23.0	432	8	ACH19886 Human adu
41	339.8	22.4	966	6	ABN74742 Bovine em
42	332.8	22.0	1001	3	AAC57452 Arachidon
43	332.8	22.0	1563	4	AAI99227 Human exc
44	332.8	22.0	1563	5	AAI63577 Human kid
45	319.4	21.1	1362	3	AAa64662 cDNA enco

ALIGNMENTS

RESULT 1
ABs62753
ID ABS62753 standard; cDNA; 1516 BP.
XX
AC ABS62753;
XX
DT 05-NOV-2002 (first entry)
XX
DE Prostate adenocarcinoma associated cDNA #28.
XX
KW Prostate cancer; differential expression; cancer diagnosis;
KW cancer treatment; cancer monitoring; prostate adenocarcinoma; gene; ss.
XX Homo sapiens.
XX US2002119463-A1.
XX 29-AUG-2002.
XX
XX 30-JUL-2001; 2001US-00919172.
XX 28-JUL-2000; 2000US-0222469P.
XX
XX (FARI/) FARIS M.
XX (TURN/) TURNER C M.
XX Faris M, Turner CM;
XX WPI: 2002-608155/65.
XX P-PSDB; ABG77176.
XX
XX New composition, useful for treating and diagnosing prostate cancer,
XX comprises human cDNAs that are differentially expressed in prostate
XX cancer.
XX
XX Claim 1; Page 66-67; 147pp; English.
XX
XX The invention describes a composition of a number of human cDNAs that are
XX differentially expressed in prostate cancer. The composition of the
XX invention useful for a high-throughput method for detecting differential
XX expression of cDNAs in a nucleic acid containing sample comprising
XX amplifying the nucleic acids of the sample, hybridizing the composition
XX with nucleic acids of the sample, detecting the hybridisation complexes
XX and comparing the complexes with those of a standard, where differences
XX indicates differential expression. The sample is from a subject with
XX prostate cancer and comparison with a standard defines an early, mid or

late stage of the disease. The composition and or protein encoded by the cDNA's are useful for a high-throughput method to screen molecules or compounds to identify a ligand that specifically binds a cDNA or protein comprising combining the composition or protein and detecting specific binding between cDNA/protein and molecule or compound. The molecules and compounds are selected from DNA, RNA, peptide nucleic acid molecules, mimetics, peptides, proteins, (ant)agonists, antibodies, immunoglobulins, inhibitors, drug compounds, pharmaceutical agents, transcription factors, repressors, and regulatory proteins. The composition is useful for diagnosing, treating or monitoring the progression or treatment of prostate cancer. The antibodies are also useful for the diagnosis of disease. This sequence represents a prostate adenocarcinoma cDNA

XX SQ Sequence 1516 BP; 493 A; 301 C; 322 G; 399 T; 0 U; 1 Other;

Query Match 99.9%; Score 1515; DB 6; Length 1516;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTGTTTGGACATAGCTGAGCCATGTACTTCAACGAGGAGCCCAATTACTTAAT 60
 DB 1 CTTTGTTTGGACATAGCTGAGCCATGTACTTCAACGAGGAGCCCAATTACTTAAT 60

QY 61 TCTGGTTGCTAGGTGGCTTCCCTTAAATCCCTATAAAATCAGAGCCCAAGTCTCCAC 120
 DB 61 TCTGGTTGCTAGGTGGCTTCCCTTAAATCCCTATAAAATCAGAGCCCAAGTCTCCAC 120

QY 121 TGCAGTGTGAATCTTCAGAGAGATTTCTTTAGTTCTTTGCAAGAGGTAGAGAT 180
 DB 121 TGCAGTGTGAATCTTCAGAGAGATTTCTTTAGTTCTTTGCAAGAGGTAGAGAT 180

QY 181 AAAGACACTTTTCAAAATGGCAATGGTATCAGAAATTCCTCAAGCAGGCTGGTTTAT 240
 DB 181 AAAGACACTTTTCAAAATGGCAATGGTATCAGAAATTCCTCAAGCAGGCTGGTTTAT 240

QY 241 GAAATGAAGAGAGGAAATATGTTCAACTGTGAAGTCATCCAAAGTGTCCCGATCA 300
 DB 241 GAAATGAAGAGAGGAAATATGTTCAAACTGTGAAGTCATCCAAAGTGTCCCGATCA 300

QY 301 GCGGTGAGCCCTATCCTACCTCAATCCATCCCTCGGATGTCGCTTGCATAAGGCC 360
 DB 301 GCGGTGAGCCCTATCCTACCTCAATCCATCCCTCGGATGTCGCTTGCATAAGGCC 360

QY 361 ATAATGGTTAAGGTGGATGAAGCAACCATATTGACATTTCTAATAGCGAACAAT 420
 DB 361 ATAATGGTTAAGGTGGATGAAGCAACCATATTGACATTTCTAATAGCGAACAAT 420

QY 421 GCACAGCTCAACAGATCAAGAGCAGCATATCTCCAGGAAACAGGAAAGCCCTGGATGA 480
 DB 421 GCACAGCTCAACAGATCAAGAGCAGCATATCTCCAGGAAACAGGAAAGCCCTGGATGA 480

QY 481 ACACGTGAAGAGCCCTTACAGGTCACTTGGAGAGGTTGTTTGTCTGCTTAATAACT 540
 DB 481 ACACGTGAAGAGCCCTTACAGGTCACTTGGAGAGGTTGTTTGTCTGCTTAATAACT 540

QY 541 CCAGGCAATTTGATGCTGATGAATCTGCTGCTGATGAAGGCGCTTGGAACTGATGAA 600
 DB 541 CCAGGCAATTTGATGCTGATGAATCTGCTGCTGATGAAGGCGCTTGGAACTGATGAA 600

QY 601 GATACCTCTAATTTGATTTGGATCAAGAACTAACAAAGAAATCAGAGACATTAACAGG 660
 DB 601 GATACCTCTAATTTGATTTGGATCAAGAACTAACAAAGAAATCAGAGACATTAACAGG 660

QY 661 GTCTACAGAGGAACTGAAGAGATCTGGCCAAAGACATAAATCCTCAGACATCTGGA 720
 DB 661 GTCTACAGAGGAACTGAAGAGATCTGGCCAAAGACATAAATCCTCAGACATCTGGA 720

QY 721 GATTTTCGGAACGCTTTGCTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 780
 DB 721 GATTTTCGGAACGCTTTGCTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 780

QY 781 AATGAAGACTTGGCTGATTCAGATGCCAGGCTTGTATGAGCAGGAGAGGAGAAAG 840

781 AATGAAGACTTGGCTGATTCAGATGCCAGGCGCTTGTATGAGCAGGAGAGAGAAAG 840

QY 841 GGGACAGACGTAACAGTGTTCATATACATCTTACACACAGAGAGCTTATCCACACTTCG 900

DB 841 GGGACAGACGTAACAGTGTTCATATACATCTTACACACAGAGAGCTTATCCACACTTCG 900

QY 901 AGAGTGTTCAGAAATACACCAAGTACAGTAAGCATGACATGAACAAAGTCTTGGACCTG 960

DB 901 AGAGTGTTCAGAAATACACCAAGTACAGTAAGCATGACATGAACAAAGTCTTGGACCTG 960

QY 961 GAGTTGAAAAGTGAATGAAATGCTTCCAGCTATCTGTAAGTGGCCCAAGGAAA 1020

DB 961 GAGTTGAAAAGTGAATGAAATGCTTCCAGCTATCTGTAAGTGGCCCAAGGAAA 1020

QY 1021 CCAGCTTTCTTTGAGAGAGAGCTTCATCAAGCCATGAAGAGTGTGGAATCGCCATAAG 1080

DB 1021 CCAGCTTTCTTTGAGAGAGAGCTTCATCAAGCCATGAAGAGTGTGGAATCGCCATAAG 1080

QY 1081 GCATGATCAGGATTTGGTTTCCCGTTCTGAAATTCGACATGAATGATATCAAGGATTC 1140

DB 1081 GCATGATCAGGATTTGGTTTCCCGTTCTGAAATTCGACATGAATGATATCAAGGATTC 1140

QY 1141 TATCAGAGAGATGATGTATCTCCCTTTGCCAGCCATCTCGATGAACCAAGAGGAGAG 1200

DB 1141 TATCAGAGAGATGATGTATCTCCCTTTGCCAGCCATCTCGATGAACCAAGAGGAGAG 1200

QY 1201 TATGAGAAAATCCCTGGTGGCTTTTGGAGGAAACTAAACATTCCTTGAATGCTCAA 1260

DB 1201 TATGAGAAAATCCCTGGTGGCTTTTGGAGGAAACTAAACATTCCTTGAATGCTCAA 1260

QY 1261 GCTATGATCAGAGACCTTAAATATATATTTTCTATAGCTTAAATAGGAAAGTTT 1320

DB 1261 GCTATGATCAGAGACCTTAAATATATATTTTCTATAGCTTAAATAGGAAAGTTT 1320

QY 1321 CTTCAACAGGATTCAGTGTAGCTACCTACATGCTGAAATATAGCTTTAAATCATTT 1380

DB 1321 CTTCAACAGGATTCAGTGTAGCTACCTACATGCTGAAATATAGCTTTAAATCATTT 1380

QY 1381 TTATATTATTAATCTGTTATATAGATAGTAACTCCATTTTAAATAATGTTTCCCAAC 1440

DB 1381 TTATATTATTAATCTGTTATATAGATAGTAACTCCATTTTAAATAATGTTTCCCAAC 1440

QY 1441 CATAAACCCCTATACAAGTTGTTCTAGTAAACATATACATGAGAAAGATGCTATGTAGCTG 1500

DB 1441 CATAAACCCCTATACAAGTTGTTCTAGTAAACATATACATGAGAAAGATGCTATGTAGCTG 1500

QY 1501 AAAATAAAATGNCGTC 1516

DB 1501 AAAATAAAATGNCGTC 1516

RESULT 2
 ABX77579
 ID ABX77579 standard; cDNA; 1516 BP.
 XX
 AC ABX77579;
 XX
 XX 09-APR-2003 (first entry)
 XX
 DE Differentially expressed breast cancer associated cDNA #74.
 XX
 XX Breast cancer; differential gene expression; BC-cDNA;
 KW breast cancer diagnosis; breast cancer monitoring;
 KW breast cancer treatment; breast cancer staging; Gene; ss.
 XX
 OS Homo sapiens.
 XX
 XX US2002156263-A1.
 DN
 XX 24-OCT-2002.
 PD
 XX 04-OCT-2001; 2001US-00974298.
 PF
 XX

05-OCT-2000; 2000US-0238331P.

(CHEN/) CHEN H.

Chen H;

WPI; 2003-182653/18.

New cDNAs, which are differentially expressed in (metastatic) breast cancer useful for diagnosing or staging, breast cancer, or for monitoring the treatment of breast cancer in an individual.

Claim 1; SEQ ID NO 92; 30pp; English.

The invention describes a combination of cDNAs (designated BC-cDNAs), which are differentially expressed in breast cancer. The combination includes 152 cDNA sequences, or their complements. The protein encoded by any of these BC-cDNAs is useful for screening several molecules or compounds to identify at least one ligand that specifically binds the protein, producing or preparing polyclonal or monoclonal antibodies, or purifying antibodies from a sample. The antibodies, which specifically bind the protein differentially expressed in breast cancer is useful for detecting the expression of a protein in a sample. The BC-cDNAs are also useful for diagnosing, monitoring the treatment of, or staging, breast cancer. This sequence represents a differentially expressed breast cancer associated cDNA. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the US patent office at
cc seqdata.uspto.gov/sequence.html?DocID=20020156263

Sequence 1516 BP; 493 A; 301 C; 322 G; 399 T; 0 U; 1 Other;

Query Match 99.9%; Score 1515; DB 7; Length 1516;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CTTTGTGTTGACATAGCTGACCATGATGCTTCAACAGAGGCGCCATTAAGT 60
1 CTTTGTGTTGACATAGCTGACCATGATGCTTCAACAGAGGCGCCATTAAGT 60
61 TCTGTTGCTAGTGTGGTCTCTTAAATCCTATAAAATCAGAAGGCCAGTCTCCAC 120
61 TCTGTTGCTAGTGTGGTCTCTTAAATCCTATAAAATCAGAAGGCCAGTCTCCAC 120
121 TCCAGTGTGAATCTTTCAGAGAAATTTCTTTAGTTCTTTCAGAGAGGTAGAGAT 180
121 TCCAGTGTGAATCTTTCAGAGAAATTTCTTTAGTTCTTTCAGAGAGGTAGAGAT 180
181 AAAGACATTTTTCAGAAATGCAATGGTATCAGAAATCCTCAAGCAGGCTGGTTATT 240
181 AAAGACATTTTTCAGAAATGCAATGGTATCAGAAATCCTCAAGCAGGCTGGTTATT 240
241 GAAATGAAGAGCAGGAATATGTTCAAACTGTGAAGTCATCAAAAGGTGGTCCCGATCA 300
241 GAAATGAAGAGCAGGAATATGTTCAAACTGTGAAGTCATCAAAAGGTGGTCCCGATCA 300
301 GCGGTGAGCCCTTATCCTACCTCATCTCGATGCTCGCTGCTTGCATTAAGGCC 360
301 GCGGTGAGCCCTTATCCTACCTCATCTCGATGCTCGCTGCTTGCATTAAGGCC 360
361 ATAAATGTTAAAGGTGTGGATGAAGCAACCATCATTGACATTTAAAGCGAAACAAAT 420
361 ATAAATGTTAAAGGTGTGGATGAAGCAACCATCATTGACATTTAAAGCGAAACAAAT 420
421 GCACAGGTCACAGATCAAGCAGCATATCTCCAGGAACAGGAGCCCTGGATGAA 480
421 GCACAGGTCACAGATCAAGCAGCATATCTCCAGGAACAGGAGCCCTGGATGAA 480
481 ACACGTGAAGAAAGCCCTTACAGTGCACCTTGGAGAGGTTGTTTAGCTCTGTAAAACT 540
481 ACACGTGAAGAAAGCCCTTACAGTGCACCTTGGAGAGGTTGTTTAGCTCTGTAAAACT 540
541 CCAGCGCAATTTGATGCTGATGAACCTTCGTGCTGCCATGAAGGCCCTTGGAACTGATGAA 600

RESULT 3
ABA09124/c
ID ABA09124 standard; cDNA; 1517 BP.
XX
AC ABA09124;

541 CCAGGCGCAATTTGATGCTGATGAACCTTCTGCTGCCATGAAGGGCTTGGAACTGATGAA 600
601 GATACTCTAAATTTGAGATTTGGCATCAAGAACTAAACAAAGAAATCAGAGACATTAACAGG 660
601 GATACTCTAAATTTGAGATTTGGCATCAAGAACTAAACAAAGAAATCAGAGACATTAACAGG 660
661 GTCTACAGAGAGAACTGAAGAGAGATCTGGCCAAAGACATAAACCCTCAGACACATCTGGA 720
661 GTCTACAGAGAGAACTGAAGAGAGATCTGGCCAAAGACATAAACCCTCAGACACATCTGGA 720
721 GATTTTCGAAACGCTTTGCTTTCTCTTCTTAAGGGTGACCGATCTGAGGACTTTGGTGTG 780
721 GATTTTCGAAACGCTTTGCTTTCTCTTCTTAAGGGTGACCGATCTGAGGACTTTGGTGTG 780
781 AATGAAGACTTCGGCTGATTCAGATGCGAGGGCTTGTATGAAGCAGGAGAAAGGAGAAAG 840
781 AATGAAGACTTCGGCTGATTCAGATGCGAGGGCTTGTATGAAGCAGGAGAAAGGAGAAAG 840
841 GGGACAGACGTAAACGCTGTTCAATACCATCTTACCCAGAGCTATCCCAACCTTCCG 900
841 GGGACAGACGTAAACGCTGTTCAATACCATCTTACCCAGAGCTATCCCAACCTTCCG 900
901 AGAGTGTTCAGAAATACACCAAGTACAGTAAAGCATGACATGAACAAAGTTCTGGACCTG 960
901 AGAGTGTTCAGAAATACACCAAGTACAGTAAAGCATGACATGAACAAAGTTCTGGACCTG 960
961 GAGTTGAAAGGTGACATTCAGAAATGCTTCACAGCTATCGTGAAAGTGCAGCAACAA 1020
961 GAGTTGAAAGGTGACATTCAGAAATGCTTCACAGCTATCGTGAAAGTGCAGCAACAA 1020
1021 CCAGCTTTCTTCGAGAGAGCTTCATCAAGCCATCAAGAGGTGTTGGAACCTGCCATAAG 1080
1021 CCAGCTTTCTTCGAGAGAGCTTCATCAAGCCATCAAGAGGTGTTGGAACCTGCCATAAG 1080
1081 GCATTGATCAGGATTAATGTTTCCCGTCTGAAATGACATGAATGATATCAAAAGCAATTC 1140
1081 GCATTGATCAGGATTAATGTTTCCCGTCTGAAATGACATGAATGATATCAAAAGCAATTC 1140
1141 TATCAGAAGATGATGGTATCTCCCTTCCAGCCATCTCGATGAAACCAAGAGAGAG 1200
1141 TATCAGAAGATGATGGTATCTCCCTTCCAGCCATCTCGATGAAACCAAGAGAGAG 1200
1201 TATGAGAAATCCTGCTGGTCTCTTTGAGGAGAACTAAACATTCCTGATGGTCTCAA 1260
1201 TATGAGAAATCCTGCTGGTCTCTTTGAGGAGAACTAAACATTCCTGATGGTCTCAA 1260
1261 GCTATGATCAGAGACATTTAAATTAATTTTCACTCTTATTAAGCTTAAATAGGAAAGTTT 1320
1261 GCTATGATCAGAGACATTTAAATTAATTTTCACTCTTATTAAGCTTAAATAGGAAAGTTT 1320
1321 CTTCAACAGGATTTACAGTGTAGCTACCTACATGCTGAAAAATATAGCCTTTAAATCATTT 1380
1321 CTTCAACAGGATTTACAGTGTAGCTACCTACATGCTGAAAAATATAGCCTTTAAATCATTT 1380
1381 TTATTTATTAACCTCTGTATTAATGAGATAGTCCATTTTTTAAATGTTTCCCAAC 1440
1381 TTATTTATTAACCTCTGTATTAATGAGATAGTCCATTTTTTAAATGTTTCCCAAC 1440
1441 CATAAABCCCTATACAGTGTGTTCTAGTAAACAATACATGAGAAAGATGCTATGTAGCTG 1500
1441 CATAAABCCCTATACAGTGTGTTCTAGTAAACAATACATGAGAAAGATGCTATGTAGCTG 1500
1501 AAAATAAAATGTCGTC 1516
1501 AAAATAAAATGTCGTC 1516

XX 11-JAN-2002 (first entry)
 XX Human lipocortin homologue-encoding cDNA, SEQ ID NO:900.
 DE
 XX Human; cytokine; cell proliferation; cell differentiation; growth factor;
 KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
 KW inhibin; chemotaxis; chemokinesis; thrombolytic; oncogenesis;
 KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
 KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
 KW chronic inflammatory condition; proliferative retinopathy;
 KW atherosclerosis; coronary heart disease; arterial ischaemia;
 KW bone disorder; osteoporosis; vascular growth disorder;
 KW tissue regeneration; wound healing; infection; immune disorder;
 KW cell culture; drug screening; gene therapy; anti-inflammatory;
 KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
 KW cytosstatic; osteopathic; haemotropic; cardiant; virucide; antibacterial;
 KW antifungal; vulnery; antiulcer; ss.
 XX Homo sapiens.
 OS
 XX WO200157188-A2.
 PN
 XX 09-AUG-2001.
 PD
 XX 05-FEB-2001; 2001WO-US003800.
 PF
 XX 03-FEB-2000; 2000US-00496814.
 PR
 XX 27-APR-2000; 2000US-00560875.
 PA
 XX (HYSE-) HYSEQ INC.
 PA
 XX Tang YT, Liu C, Dmanac RT;
 PI
 XX WPI; 2001-457740/49.
 DR
 XX P-PSDB; ABB11880.
 XX
 Human proteins and DNA encoding sequences useful for preventing, treating
 PT or ameliorating a medical condition in a mammalian subject e.g. arthritis
 PT and cancer.
 PT
 PS Claim 1; Page 783; 1963pp; English.
 XX
 Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
 CC sequences ABB08225-ABB09574 represent nucleic acids encoding them. The
 CC invention also relates to vectors and recombinant host cells comprising a
 CC nucleotide of the invention, methods of producing the novel polypeptides,
 CC antibodies against the polypeptides, methods of detecting the nucleotides
 CC or polypeptides in a sample, and methods of identifying compounds which
 CC bind to polypeptides of the invention. Although novel, many of the
 CC polypeptides of the invention have homology to known proteins, thereby
 CC giving an insight into their probable biological activities, and hence
 CC potential therapeutic applications. The polypeptides of the invention may
 CC have various activities, including cytokine, cell proliferation or cell
 CC differentiation activities; stem cell growth factor activity;
 CC haematopoiesis regulatory activity; tissue growth activity;
 CC immunomodulatory activity; activin- or inhibin-related activities;
 CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
 CC thrombolytic activities; receptor or ligand activities; or may be
 CC involved in oncogenesis, cancer cell proliferation or metastasis.
 CC Depending on their biological activities, polypeptides and nucleotides of
 CC the invention are useful for preventing, treating or ameliorating medical
 CC conditions, e.g., by protein or gene therapy. Such conditions include
 CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
 CC vascular growth. Polypeptides involved with tissue regeneration and
 CC repair (or nucleic acids encoding them) may be used to promote wound
 CC healing (e.g., of burns, incisions and ulcers), while those with
 CC immunomodulatory activities may be used in the treatment of viral,
 CC bacterial and fungal infections in addition to immune disorders.
 CC Polypeptides with growth factor activity may be used in cell cultures to

CC promote cell growth. For example, such polypeptides may be used to
 CC manipulate stem cells in culture to give rise to neuroepithelial cells
 CC that can be used to augment or replace cells damaged by illness,
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides
 CC may also be used in the diagnosis of the above conditions, and in drug
 CC screening techniques. The present sequence represents a cDNA encoding a
 CC novel human polypeptide of the invention
 XX
 SQ Sequence 1517 BP; 391 A; 319 C; 300 G; 507 T; 0 U; 0 Other;
 Query Match 98.8%; Score 1497.4; DB 4; Length 1517;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1498; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 17 AGCTCAGCCAGTACTTCAACACAGAGCGCAGCCAAATTAACCTTCTGGTTCCTAGGTGT 76
 DB 1517 AGCTGAGCAGTACTTCAACACAGAGCGCAGCCAAATTAACCTTCTGGTTCCTAGGTGT 1458
 QY 77 GGCCTTCCTTTAAATTCCTATMAAATCAGAGCCCAAGTCTCCACTGCCAGTGTGAATCT 136
 DB 1457 GGCCTTCCTTTAAATTCCTATMAAATCAGAGCCCAAGTCTCCACTGCCAGTGTGAATCT 1398
 QY 137 TCAGAGAGAGATTTCTCTTTAGTTCTTTTGCAGAGGTAGAGATAAAGACACTTTTCAA 196
 DB 1397 TCAGAGAGAGATTTCTCTTTAGTTCTTTTGCAGAGGTAGAGATAAAGACACTTTTCAA 1338
 QY 197 AAATGCCAATGGTATCAGAAATTCCTCAAGCAGGCTGGTTTATTGAAATGAAGAGAGG 256
 DB 1337 AAATGCCAATGGTATCAGAAATTCCTCAAGCAGGCTGGTTTATTGAAATGAAGAGAGG 1278
 QY 257 AATATGTTCAAACTGTGAAGTCAATCCAAAGGTGTCCCGGATCAGCGTGAGCCCTATC 316
 DB 1277 AATATGTTCAAACTGTGAAGTCAATCCAAAGGTGTCCCGGATCAGCGTGAGCCCTATC 1218
 QY 317 CTACCTTCAATCCATCCTCGGATGTGCTGCTTGCATTAAGGCCCATTAATGGTTAAAGGTG 376
 DB 1217 CTACCTTCAATCCATCCTCGGATGTGCTGCTTGCATTAAGGCCCATTAATGGTTAAAGGTG 1158
 QY 377 TGGATGAAGCAACCATTAATGACATTTCTAATAGCGAACAATGCGAGCGTGCAACAGA 436
 DB 1157 TGGATGAAGCAACCATTAATGACATTTCTAATAGCGAACAATGCGAGCGTGCAACAGA 1098
 QY 437 TCAAAAGCAGCATATCTCCAGGAAACAGGAAAGCCCTGGATGAAACACTGAAAGAGGCC 496
 DB 1097 TCAAAAGCAGCATATCTCCAGGAAACAGGAAAGCCCTGGATGAAACACTGAAAGAGGCC 1038
 QY 497 TTACAGGTCACTTCCAGAGGTGTGTTTGTCTGTCTTAAATCCAGCGCAATTTGATG 556
 DB 1037 TTACAGGTCACTTCCAGAGGTGTGTTTGTCTGTCTTAAATCCAGCGCAATTTGATG 978
 QY 557 CTGATGAATCTGCTGCTGCTGATGAAGGGCTTGGAACTGTGAAGATACTCTTAATTGAGA 616
 DB 977 CTGATGAATCTGCTGCTGCTGATGAAGGGCTTGGAACTGTGAAGATACTCTTAATTGAGA 918
 QY 617 TTTTGGCATCAAGAACTAACAAAGAAATCAGAGACATTAACAGGGTCTACAGAGAGGAAC 676
 DB 917 TTTTGGCATCAAGAACTAACAAAGAAATCAGAGACATTAACAGGGTCTACAGAGAGGAAC 858
 QY 677 TGAAGAGAGATCTGCCAAAGACATAAAGCTCAGACACATCTGGAGATTTTCGAAACGGTT 736
 DB 857 TGAAGAGAGATCTGCCAAAGACATAAAGCTCAGACACATCTGGAGATTTTCGAAACGGTT 798
 QY 737 TGGTTTCTCTTCTTAAAGGTGTCCGATCTGAGGACTTTGGTGTGTAATGAAGACTTTGGGTG 796
 DB 797 TGGTTTCTCTTCTTAAAGGTGTCCGATCTGAGGACTTTGGTGTGTAATGAAGACTTTGGGTG 738
 QY 797 ATTACAGTCCAGGGCTTGTATGAAGCAGGAGAGAGAGAAAGGGGACAGACGCTAAACG 856
 DB 737 ATTACAGTCCAGGGCTTGTATGAAGCAGGAGAGAGAGAAAGGGGACAGACGCTAAACG 678
 QY 857 TGTTCATATACCATCTCTTACCAACAGAGCTATCCCAACTTCGCGAGAGTGTTCGAGAAAT 916
 DB 677 TGTTCATATACCATCTCTTACCAACAGAGCTATCCCAACTTCGCGAGAGTGTTCGAGAAAT 618

QY	917	ACACCAAGTACAGTAAGCATGA	CATGAACAAAGTTCTGGACCTTGGAGTTGA	AAAGGTGACA	976
Db	617	ACACCAAGTACAGTAAGCATGA	CATGAACAAAGTTCTGGACCTTGGAGTTGA	AAAGGTGACA	558
QY	977	TTGAGAAATGSCCTCAGAGCT	TATCGTGAAGTGCGCCAACGACAAACAGCTTTCTTTGCGAG	1036	
Db	557	TTGAGAAATGSCCTCAGAGCT	TATCGTGAAGTGCGCCAACGACAAACAGCTTTCTTTGCGAG	498	
QY	1037	AGAAGCTTCATCAAGGCCAT	GAAGGTGTTGGAACTCGCCATAAGGCCATTGATCAGAGTTA	1096	
Db	497	AGAAGCTTCATCAAGGCCAT	GAAGGTGTTGGAACTCGCCATAAGGCCATTGATCAGAGTTA	438	
QY	1097	TGGTTTCCCCTCTGGAATTC	GATGATCATCAAGCATTCATCAAGAGCATTCATCAAGAGATGTATG	1156	
Db	437	TGGTTTCCCCTCTGGAATTC	GATGATCATCAAGCATTCATCAAGAGCATTCATCAAGAGATGTATG	378	
QY	1157	GTATCTCCCTTTGCCAAGCC	ATCCTGGATGAAACCAAAGGAGAGTAGTAGGAAATCTCTGG	1216	
Db	377	GTATCTCCCTTTGCCAAGCC	ATCCTGGATGAAACCAAAGGAGAGTAGTAGGAAATCTCTGG	318	
QY	1217	TGGCTCTTTCTGGAGGAA	CTAAACATTCCTCTGGATCGTCTCAAGCTATGATCAGAGAC	1276	
Db	317	TGGCTCTTTCTGGAGGAA	CTAAACATTCCTCTGGATCGTCTCAAGCTATGATCAGAGAC	258	
QY	1277	TTTAAATATATATTTTCAT	CTCTATAAGCTTAAATAGGAAAGTTTCTTTCAACAGGATTACA	1336	
Db	257	TTTAAATATATATTTTCAT	CTCTATAAGCTTAAATAGGAAAGTTTCTTTCAACAGGATTACA	198	
QY	1337	GTGTAGTCACTACATGCT	GAAATAATAGCCTTTAAATCATTTTATATATATATTA	1396	
Db	197	GTGTAGTCACTACATGCT	GAAATAATAGCCTTTAAATCATTTTATATATATATTA	138	
QY	1397	TATAATAGAGATAAGTCC	ATTTTTTAAAAATGTTTTCCCCAAACCATAAAAACCTTATACA	1456	
Db	137	TATAATAGAGATAAGTCC	ATTTTTTAAAAATGTTTTCCCCAAACCATAAAAACCTTATACA	78	
QY	1457	AGTTGTTCTTAGTAACA	ATACATGAGAAGATGTCTATGTAGCTGAAATATAAATGTCGTC	1516	
Db	77	AGTTGTTCTTAGTAACA	ATACATGAGAAGATGTCTATGTAGCTGAAATATAAATGTCGTC	18	

RESULT 4	
AAH72822	
ID	AAH72822 standard; cDNA; 2334 BP.
XX	
XX	AAH72822;
XX	AC
XX	AC
XX	19-SEP-2001 (first entry)
DT	
XX	Human cervical cancer marker nucleic acid 4096.
DE	
XX	Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
KW	
XX	
OS	Homo sapiens.
XX	
FN	WO200142467-A2.
XX	
XX	14-JUN-2001.
PD	
XX	
PF	08-DEC-2000; 2000WO-US033312.
XX	
PR	08-DEC-1999; 99US-0169681P.
PR	21-DEC-1999; 99US-0171350P.
PR	14-MAR-2000; 2000US-0189315P.
PR	12-MAY-2000; 2000US-0203791P.
PR	09-JUN-2000; 2000US-0210600P.
PR	21-JUL-2000; 2000US-0220114P.
XX	
PA	(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX	
PI	Schlegel R, Deeds J, Berger A, Zhao X;
XX	

DR WPI: 2001-375006/39.

XX New isolated nucleic acid for diagnosing and treating cervical cancer and
PT for assessing and detecting compounds for treating the cancer.

XX

XX Claim 1; Page 870-871; 1051pp; English.

PS

XX The invention relates to novel genes (AAH68727-AAH73383) associated with
CC cervical cancer with cytostatic activity. The nucleic acids and encoded
CC polypeptides are useful: to assess if a patient is afflicted with
CC cervical cancer or has a pre-malignant condition; to monitor the
CC progression of cervical cancer or a premalignant condition in a patient;
CC and to select and/or assess the efficacy of a compound or therapy for
CC inhibiting cervical cancer in a patient. The nucleic acids may also be
CC useful for gene therapy

XX

SQ Sequence 2334 BP; 774 A; 570 C; 510 G; 480 T; 0 U; 0 Other;

Query Match 98.7%; Score 1495.6; DB 4; Length 2334;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1498; Conservative 0; Mismatches 5; Indels 0; Gaps 0

QY 14 CATAGCTCAGGCATGTACTTCAACAGAGGCGACCAATTACTAATTCTGTGTTGCTAGG 73
Db 16 CGTGGTGCAGCATGTACTTCAACAGAGGCGACCAATTACTAATTCTGTGTTGCTAGG 75

QY 74 TGTGCTTCCTTTAAATCCTATAAATCAGAGCCCAAGTCTCCACTGCCAGTGTGA 133
Db 76 TGTGCTTCCTTTAAATCCTATAAATCAGAGCCCAAGTCTCCACTGCCAGTGTGA 135

QY 134 TCTTCAGAGAAGAATTTCTTTTAGTTTTCCTGCAAGAAGTAGAGATAAGACACTTTTT 193
Db 136 TCTTCAGAGAAGAATTTCTTTTAGTTTTCCTGCAAGAAGTAGAGATAAGACACTTTTT 195

QY 194 CAAAAATGGCAATGGTATCAGAAATTCCTCAAGCAGCGCTGGTTTATTGAAATGAAGAGC 253
Db 196 CAAAAATGGCAATGGTATCAGAAATTCCTCAAGCAGCGCTGGTTTATTGAAATGAAGAGC 255

QY 254 AGGAATATGTTCAAACTGTGAAGTCATCCAAAGGTGTCCTCGGATCAGCGGTGAGCCCT 313
Db 256 AGGAATATGTTCAAACTGTGAAGTCATCCAAAGGTGTCCTCGGATCAGCGGTGAGCCCT 315

QY 314 ATCTACCTTCAATCCATCTCGATCTCGCTGCTGCAATAGGCATTAATGGTTAAAG 373
Db 316 ATCTACCTTCAATCCATCTCGATCTCGCTGCTGCAATAGGCATTAATGGTTAAAG 375

QY 374 GTGTGGATGAAGCAACCATTAATTGACATTTCTAACTAAGCGAAAACAATGCACAGCGTCAAC 433
Db 376 GTGTGGATGAAGCAACCATTAATTGACATTTCTAACTAAGCGAAAACAATGCACAGCGTCAAC 435

QY 434 AGATCAAAAGCAGCATATCTCCAGAAAACAGGAAGCCCTGGATGAAACACTCAAGAAAAG 493
Db 436 AGATCAAAAGCAGCATATCTCCAGAAAACAGGAAGCCCTGGATGAAACACTCAAGAAAAG 495

QY 494 CCCTTACAGGTCACTTGCAGGAGTGTGTTTAGCTTGTCTTAAATACTCCAGCGCAATTTG 553
Db 496 CCCTTACAGGTCACTTGCAGGAGTGTGTTTAGCTTGTCTTAAATACTCCAGCGCAATTTG 555

QY 554 ATGCTGATGAATTCGTGCTGCCATGAGGCGCTTGGAACCTGATGAAGATPACTCTAATTG 613
Db 556 ATGCTGATGAATTCGTGCTGCCATGAGGCGCTTGGAACCTGATGAAGATPACTCTAATTG 615

QY 614 AGATTTTGGCATCAAGAACTAACAAAGAAATCAGAGACATTAACAGGCTCTACAGAGAG 673
Db 616 AGATTTTGGCATCAAGAACTAACAAAGAAATCAGAGACATTAACAGGCTCTACAGAGAG 675

QY 674 AACTGAAGAGCATCTGGCCCAAGACATAACCTCAGACACATCTGGAGATTTTCGGAACG 733
Db 676 AACTGAAGAGCATCTGGCCCAAGACATAACCTCAGACACATCTGGAGATTTTCGGAACG 735

QY 734 CTTTGTCTTTCTTCTTGAAGGGTACCGAATCTGAGGACTTTTGTGTGAATGAAGACTTGG 793
Db 736 CTTTGTCTTTCTTCTTGAAGGGTACCGAATCTGAGGACTTTTGTGTGAATGAAGACTTGG 795

[illegible]

RESULT 5	
AAH72950	
ID	AAH72950 standard; cDNA; 2334 BP.
XX	
XX	AAH72950;
XX	
XX	19-SEP-2001 (first entry)
DT	
XX	
XX	Human cervical cancer marker nucleic acid 4224.
DE	
XX	
XX	Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
KW	
XX	
XX	Homo sapiens.
OS	
XX	
XX	WO200142467-A2.
FN	
XX	
PD	14-JUN-2001.
XX	
XX	08-DEC-2000; 2000WO-US033312.
XX	

```

XX      08-DEC-1999;    99US-0169681P.
PR      21-DEC-1999;    99US-0171350P.
PR      14-MAR-2000;    2000US-0189315P.
PR      12-MAY-2000;    2000US-0203791P.
PR      09-JUN-2000;    2000US-0210600P.
PR      21-JUL-2000;    2000US-0220114P.
XX      (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX      Schlegel R, Deeds J, Berger A, Zhao X;
PI      WPI; 2001-375006/39.
DR
XX      New isolated nucleic acid for diagnosing and treating cervical cancer and
PT      for assessing and detecting compounds for treating the cancer.
PE
XX      Claim 1; Page 925; 105lpp; English.
PS
XX      The invention relates to novel genes (AHU69727-AAH73393) associated with
CC      cervical cancer with cytostatic activity. The nucleic acids and encoded
CC      polypeptides are useful: to assess if a patient is afflicted with
CC      cervical cancer or has a pre-malignant condition; to monitor the
CC      progression of cervical cancer or a premalignant condition in a patient;
CC      and to select and/or assess the efficacy of a compound or therapy for
CC      inhibiting cervical cancer in a patient. The nucleic acids may also be
CC      useful for gene therapy
XX
XX      Sequence 2334 BP; 774 A; 570 C; 510 G; 480 T; 0 U; 0 Other;
SQ
Query Match          98.7%; Score 1495.6; DB 4; Length 2334;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1498; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      14 CATAGCTGAGCGCATGTACTTCAACACGAGGGCAGGCCAAATTACTTAATTTCTGTTCCTTAGG 73
DB      16 CGTGCGTGAGGCATGTACTTCAACACGAGGGCAGGCCAATTACTTAATTTCTGTTCCTTAGG 75

QY      74 TGTCGCTTCCTTTAAAAATCCTATAAAAATCAGAAGCCCAAAGTCTCCAAGTCCCAAGTGTGA 133
DB      76 TGTCGCTTCCTTTAAAAATCCTATAAAAATCAGAAGCCCAAAGTCTCCAAGTCCCAAGTGTGA 135

QY      134 TCCTTCAGAGAGAAATTTCTCTTTAGTTCTTTGCAAGAAGGTAGAGATAAAGACACTTTTT 193
DB      136 TCCTTCAGAGAGAAATTTCTCTTTAGTTCTTTGCAAGAAGGTAGAGATAAAGACACTTTTT 195

QY      194 CAAAAATGGCAATGGTATCAGAAATTCCTCAAGCAGGCCCTGGTTTATTGAAATAGAGAGC 253
DB      196 CAAAAATGGCAATGGTATCAGAAATTCCTCAAGCAGGCCCTGGTTTATTGAAATAGAGAGC 255

QY      254 AGGATATGTTCAACATGTGAAGTCATCCAAAGAGTGTCCCGGATCAGCGGTGAGCCCT 313
DB      256 AGGATATGTTCAACATGTGAAGTCATCCAAAGAGTGTCCCGGATCAGCGGTGAGCCCT 315

QY      314 ATCTACCCTTCAATCCATCCTCCGATTCGCTGCCTTGGATAGGCCATAATGGTTTAAAG 373
DB      316 ATCTACCCTTCAATCCATCCTCCGATTCGCTGCCTTGGATAGGCCATAATGGTTTAAAG 375

QY      374 GTGTGGATGAGCACCATCATTCACATTTCTAAGCGAAACAATGCACAGCGTCAAC 433
DB      376 GTGTGGATGAGCACCATCATTCACATTTCTAAGCGAAACAATGCACAGCGTCAAC 435

QY      434 AGATCAAAGCAGCATATCTCCAGAAAACAGGAAAGCCCCTGGATGAAAACACTGAAGAAG 493
DB      436 AGATCAAAGCAGCATATCTCCAGAAAACAGGAAAGCCCCTGGATGAAAACACTGAAGAAG 495

QY      494 CCCCTACAGGTCACTTCGAGAGGTTGTTTTAGCTCTGCTFAAAACTCCAGCGCAATTG 553
DB      496 CCCCTACAGGTCACTTCGAGAGGTTGTTTTAGCTCTGCTFAAAACTCCAGCGCAATTG 555

QY      554 ATGCTGATGAACCTTCGTCTGCCATGAAGGCCCTTGGAACTGATGAAGATATCTTAATTG 613
DB      556 ATGCTGATGAACCTTCGTCTGCCATGAAGGCCCTTGGAACTGATGAAGATATCTTAATTG 615

```


Db 241 TGGTTAAAGGTGTGGATGAAGCAACCATTCATTGACATTTCTAACTAAGCGAAACAATGTCAC 300
 QY 425 AGCGTCAACAGATCAAGCAGCATATCTCCAGAAACAGAAAGCCCTGGATGAACAC 484
 Db 301 AGCGTCAACAGATCAAGCAGCATATCTCCAGAAACAGAAAGCCCTGGATGAACAC 360
 QY 485 TGAAGAAAGCCCTTACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 544
 Db 361 TGAAGAAAGCCCTTACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 420
 QY 545 CGCAATTTGATGTGATGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 604
 Db 421 CGCAATTTGATGTGATGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
 QY 605 CTCTAATTTGATGTGATGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 664
 Db 481 CTCTAATTTGATGTGATGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
 QY 665 ACAGAGAGGAACCTGAGAGAGATCTGCGCAAGACATTAACCTCAGACACATCTGGAGAT 724
 Db 541 ACAGAGAGGAACCTGAGAGAGATCTGCGCAAGACATTAACCTCAGACACATCTGGAGAT 500
 QY 725 TTCGGAACCGTTTCT 784
 Db 601 TTCGGAACCGTTTCT 660
 QY 785 AAGACTTGGCTGATTCAGATGCGAGGCGCTTGTATGAAGCAGGAGAAAGGAGGGA 844
 Db 661 AAGACTTGGCTGATTCAGATGCGAGGCGCTTGTATGAAGCAGGAGAAAGGAGGGA 720
 QY 845 CAGACGTAACCGTGTCTCAATACCATCTTACCACCAAGATCTCCACCACTTCGCAGAG 904
 Db 721 CAGACGTAACCGTGTCTCAATACCATCTTACCACCAAGATCTCCACCACTTCGCAGAG 780
 QY 905 TGTTTCAGAAATACCAAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAG 964
 Db 781 TGTTTCAGAAATACCAAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAG 840
 QY 965 TGAAGGTGACATTTGAGAAATGCTCTACAGCTATCGTGAAGTGGCGCCACCAAGCAACCCAG 1024
 Db 841 TGAAGGTGACATTTGAGAAATGCTCTACAGCTATCGTGAAGTGGCGCCACCAAGCAACCCAG 900
 QY 1025 CTTTCTTTGAGAGAGGTTTCATCAGCCATGAAGGTGTTGAAGTGGCGCCATTAAGGCAT 1084
 Db 901 CTTTCTTTGAGAGAGGTTTCATCAGCCATGAAGGTGTTGAAGTGGCGCCATTAAGGCAT 960
 QY 1085 TGATCAGGATTTATGTTTCCGTTCTGAAATTCGATGATGATGATGATGATGATGATGATG 1144
 Db 961 TGATCAGGATTTATGTTTCCGTTCTGAAATTCGATGATGATGATGATGATGATGATG 1020
 QY 1145 AGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1204
 Db 1021 AGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
 QY 1205 AGAAATCTCGTGGCTCTTTGCGAGGAACTAAACATTCCTTGTGATGCTCAAGCTA 1264
 Db 1081 AGAAATCTCGTGGCTCTTTGCGAGGAACTAAACATTCCTTGTGATGCTCAAGCTA 1140
 QY 1265 TGATCAGAGACTTTAAT 1324
 Db 1141 TGATCAGAGACTTTAAT 1200
 QY 1325 AACAGGATTCAGTGTAGCTACCTACCTGCTGAAATATATAGCCCTTTAAATCATTTTAT 1384
 Db 1201 AACAGGATTCAGTGTAGCTACCTACCTGCTGAAATATATAGCCCTTTAAATCATTTTAT 1260
 QY 1385 ATTATACTCTGTATAATAGAGATAGTCCATTTTAAAAATGTTTCCCAACCAATA 1444
 Db 1261 ATTATACTCTGTATAATAGAGATAGTCCATTTTAAAAATGTTTCCCAACCAATA 1320
 QY 1445 AAACCCCTATACAGTGTGTTCTAGTAAACATACATGAGAAAGATGCTATGTAGCTGAAA 1504

Db 1321 AAACCCCTATACAGTGTGTTCTTAGTAAACATACATGAGAAAGATGTTCTATGTAGCTGAAA 1380
 QY 1505 TAAATGNCCTC 1516
 Db 1381 TAAATGACGTC 1392
 RESULT 7
 AAX57357
 ID AAX57357 standard; DNA; 1399 BP.
 XX AAX57357;
 AC AAX57357;
 DT 26-JUL-1999 (first entry)
 XX Human p-40/annexin I DNA.
 DE P-40; annexin I; human; MDR; multidrug resistance; antisense; prevention;
 KW cancer treatment; detection; pathogen; therapeutic agent;
 KW antifungal agent; plant; resistance; antitumor; antifungal; ss.
 OS Homo sapiens.
 XX WO9921980-A1.
 PN 06-MAY-1999.
 XX 26-OCT-1998; 98WO-CA000992.
 PF 24-OCT-1997; 97CA-02219299.
 PR (UYNM-) UNIV MCGILL.
 FA Georges E, Wang Y;
 PI WPI: 1999-337419/28.
 DR P-PSDB; AAY08412.
 XX Modulating or assessing multidrug resistance related to annexin proteins.
 PT Claim 5; Fig 1B-C; 63pp; English.
 PS This invention describes a novel human annexin family member, P-40 (also
 CC known as annexin I) which is a member of the MDR (multidrug resistance)
 CC gene family, for assessing or modulating MDR in a cell. Antisense P-40
 CC sequences are used to prevent MDR in animals, particularly in conjunction
 CC with cancer treatment. Detecting levels of the P-40 nucleic acid, or
 CC related RNA, is used to detect cancer (or pathogens) with MDR, or
 CC susceptibility. P-40 nucleic acid can also be used as a target for
 CC identifying therapeutic agents, e.g. antifungal agents, and increasing
 CC the nucleic acid expression in plants may be used to develop specific
 CC resistance. The products of the invention have antitumor and antifungal
 CC activity
 XX Sequence 1399 BP; 464 A; 273 C; 298 G; 364 T; 0 U; 0 Other;
 SQ Query Match 91.5%; Score 1387.8; DB 2; Length 1399;
 Best Local Similarity 99.8%; Pred. NO. 0;
 Matches 1399; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 125 AGTGTGAATCTTCAGAGAAGATTTCTTTAGTCTTTTCAGAGAAGGTAGAGATAAG 184
 Db 1 AGTGTGAATCTTCAGAGAAGATTTCTTTAGTCTTTTCAGAGAAGGTAGAGATAAG 60
 QY 185 ACACATTTTCAAAAATGCAATGGTATCAGATTCCTCAAGCAGGCGCTTTATTGAAA 244
 Db 61 ACATTTTCAAAAATGCAATGGTATCAGATTCCTCAAGCAGGCGCTTTATTGAAA 120
 QY 245 ATGAAGAGCAGGAATATGTTCAAACCTGTAAGTCTCAAAGGTGGTCCCGATCAGCGG 304
 Db 121 ATGAAGAGCAGGAATATGTTCAAACCTGTAAGTCTCAAAGGTGGTCCCGATCAGCGG 180
 QY 305 TGAGCCCTTCTCTACCTTCAATTCATCTCGATGTCGCTTGCCTTGCATAGGCCATAA 364

181 TGAGCCCTTCTCTACCTTCAATCCATCTCGATGTGCTGCTTGTGATAGGCCATAA 240
365 TGGTTAAAGGTGTGATGAAGCAACATCATTTGACATTTCTAATGAAGCGAAACAATGCAC 424
241 TGGTTAAAGGTGTGATGAAGCAACATCATTTGACATTTCTAATGAAGCGAAACAATGCAC 300
425 AGCGTCAACAGATCAAGCAGCATATCTCCAGGAACAGGAAGCCCTGGATGAACAC 484
301 AGCGTCAACAGATCAAGCAGCATATCTCCAGGAACAGGAAGCCCTGGATGAACAC 360
485 TGAAGAAAGCCCTTACAGCTCACTTGAGAGGTTGTTTGTAGCTCTGCTTAAAACTCCAG 544
361 TTAAGAAAGCCCTTACAGCTCACTTGAGAGGTTGTTTGTAGCTCTGCTTAAAACTCCAG 420
545 CGCAATTTGATCTGATGAACCTCGTCTGCCATGAAGGGCCCTTGAATGATGAAGATA 604
421 CGCAATTTGATCTGATGAACCTCGTCTGCCATGAAGGGCCCTTGAATGATGAAGATA 480
605 CTCTAATTGAGATTTTGGCATCAAGAACTTAACAAAGAAATCAGACATTAACAGGGTCT 664
481 CTCTAATTGAGATTTTGGCATCAAGAACTTAACAAAGAAATCAGACATTAACAGGGTCT 540
665 ACAGAGAGAACTGAGAGAGATCTGGCCAGAGACATTAACCTCAGACATCTGGAGATT 724
541 ACAGAGAGAACTGAGAGAGATCTGGCCAGAGACATTAACCTCAGACATCTGGAGATT 600
725 TTCGGAACGCTTTGCTTTCTCTTGTCTAAGGGTGAACGATCTGAGGACTTTGGTGTGAATG 784
601 TTCGGAACGCTTTGCTTTCTCTTGTCTAAGGGTGAACGATCTGAGGACTTTGGTGTGAATG 660
785 AGACTTGGCTGATTCAGATGCGAGGGCTTGTATGAAGCAGGAGAGAGAGAGAGGGA 844
661 AGACTTGGCTGATTCAGATGCGAGGGCTTGTATGAAGCAGGAGAGAGAGAGAGGGA 720
845 CAGACGTAAACGTTTCAATACCATCTTACACACAGAGAGCTATCCAACTTCGCAGAG 904
721 CAGACGTAAACGTTTCAATACCATCTTACACACAGAGCTATCCAACTTCGCAGAG 780
905 TGTTCAGAAATACACAAAGTACAGTAAGCATGACATGAACAAAGTTCTGGACCTGGAGT 964
781 TGTTCAGAAATACACAAAGTACAGTAAGCATGACATGAACAAAGTTCTGGACCTGGAGT 840
965 TGAAGGTGACATTTGAGAAATGCTTCAAGCTATCGTGAAGTGGCCACAGCAACACAG 1024
841 TGAAGGTGACATTTGAGAAATGCTTCAAGCTATCGTGAAGTGGCCACAGCAACACAG 900
1025 CTTCCTTTGAGAGAGCTTTCATCAAGCCATGAAGGTTGTGAACTCGCCATAAGGCAT 1084
901 CTTCCTTTGAGAGAGCTTTCATCAAGCCATGAAGGTTGTGAACTCGCCATAAGGCAT 960
1085 TGATCAGGATTTGGTTTCCCGTTCTGAAATGACATGAATGACATCAAGCATTTCTATC 1144
961 TGATCAGGATTTGGTTTCCCGTTCTGAAATGACATGAATGACATCAAGCATTTCTATC 1020
1145 AGAAGATGTATGTTATCTCCCTTTGGCAAGCCATCTCGATGAACCAAGAGAGATATG 1204
1021 AGAAGATGTATGTTATCTCCCTTTGGCAAGCCATCTCGATGAACCAAGAGAGATATG 1080
1205 AGAAATCTGTTGGCTCTTTTGGAGGAACCTTAACATTCCTTTCATGCTCTCAGCTA 1264
1081 AGAAATCTGTTGGCTCTTTTGGAGGAACCTTAACATTCCTTTCATGCTCTCAGCTA 1140
1265 TGATCAGAAGACTTTAATATATATTTTCACTTCAATGCTTAAGCTTAAGTAAAGTTTCTTC 1324
1141 TGATCAGAAGACTTTAATATATATTTTCACTTCAATGCTTAAGCTTAAGTAAAGTTTCTTC 1200
1325 AACAGATTACGTGTAGCTACCTACATGCTGAAATATAGCTTTAATCAATTTTAT 1384
1201 AACAGATTACGTGTAGCTACCTACATGCTGAAATATAGCTTTAATCAATTTTAT 1260
1385 ATTATAACTCTGTATATAGAGTAAGTCCATTTTAAAAATGTTTTCCCAAAACATA 1444

1261 ATTATAACTCTGTATATAGAGATAAGTCCATTTTAAAAATGTTTTCCCAAAACATA 1320
1445 AAACCCCTATACAGTTGTTCTAGTAAACAATACATGAGAAAGATGTTCTATGTAGCTGAAA 1504
1321 AAACCCCTATACAGTTGTTCTAGTAAACAATACATGAGAAAGATGTTCTATGTAGCTGAAA 1380
1505 TAAAAATGNCGT 1516
1381 TAAAAATGACGTC 1392
RESULT 8
ABL66809
ID ABL66809 standard; DNA; 1399 BP.
XX ABL66809;
XX AC ABL66809;
XX XX
XX DT 15-MAY-2002 (first entry)
XX XX
XX DE Lung cancer related gene sequence SEQ ID NO:5146.
XX XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; Gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX XX
XX OS Homo sapiens.
XX XX
XX PN WO200194629-A2.
XX XX
XX PD 13-DEC-2001.
XX XX
XX PF 30-MAY-2001; 2001WO-US010838.
XX XX
PR 05-JUN-2000; 2000US-0209473P.
PR 05-JUN-2000; 2000US-0209531P.
PR 18-SEP-2000; 2000US-023133P.
PR 18-SEP-2000; 2000US-0233617P.
PR 20-SEP-2000; 2000US-0234009P.
PR 20-SEP-2000; 2000US-0234034P.
PR 20-SEP-2000; 2000US-0234052P.
PR 22-SEP-2000; 2000US-0234509P.
PR 22-SEP-2000; 2000US-0234567P.
PR 25-SEP-2000; 2000US-0234923P.
PR 25-SEP-2000; 2000US-0234924P.
PR 25-SEP-2000; 2000US-0235077P.
PR 25-SEP-2000; 2000US-0235082P.
PR 25-SEP-2000; 2000US-0235134P.
PR 25-SEP-2000; 2000US-0235280P.
PR 26-SEP-2000; 2000US-0235637P.
PR 26-SEP-2000; 2000US-0235638P.
PR 27-SEP-2000; 2000US-0235711P.
PR 27-SEP-2000; 2000US-0235720P.
PR 27-SEP-2000; 2000US-0235840P.
PR 27-SEP-2000; 2000US-0235863P.
PR 28-SEP-2000; 2000US-0236028P.
PR 28-SEP-2000; 2000US-0236032P.
PR 28-SEP-2000; 2000US-0236033P.
PR 28-SEP-2000; 2000US-0236034P.
PR 28-SEP-2000; 2000US-0236109P.
PR 28-SEP-2000; 2000US-0236111P.
PR 29-SEP-2000; 2000US-0236842P.
PR 29-SEP-2000; 2000US-0236891P.
PR 02-OCT-2000; 2000US-0237172P.
PR 02-OCT-2000; 2000US-0237173P.
PR 02-OCT-2000; 2000US-0237278P.
PR 02-OCT-2000; 2000US-0237294P.
PR 02-OCT-2000; 2000US-0237295P.
PR 02-OCT-2000; 2000US-0237316P.
PR 03-OCT-2000; 2000US-0237435P.
PR 03-OCT-2000; 2000US-0237598P.
PR 03-OCT-2000; 2000US-0237604P.
PR 03-OCT-2000; 2000US-0237606P.

PR	03-OCT-2000; 2000US-0237608P.
PR	01-NOV-2000; 2000US-0244867P.
PR	01-NOV-2000; 2000US-0245084P.
PA	(AVAL-) AVALON PHARM.
XX	
XX	Young PE, Augustus M., Carter KC, Ebner R, Endress G, Horrigan S;
PI	Soppet DR, Weaver Z;
FI	
DR	WPI; 2002-198264/24.
XX	
XX	Screening for anti-neoplastic agent involves exposing cells to a chemical
PPT	agent to be tested for anti-neoplastic activity, and determining a change
PT	in expression of a gene of a signature gene set.
XX	
PS	Claim 1; SEQ ID NO 5146; 44pp; English.
PS	
CC	The present invention describes a method (M1) for screening for an anti-
CC	neoplastic agent. The method involves exposing cells to a chemical agent
CC	to be tested for anti-neoplastic activity, determining a change in
CC	expression of at least one gene (I) of a signature gene set, where (I)
CC	comprises a sequence (S) selected from 8447 sequences (given in ABU61664
CC	to ABU70110), or is at least 95% identical to (S), where a change in
CC	expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC	activity and can be used in gene therapy. M1 can be used for screening an
CC	anti-neoplastic agent, and can be used for producing a product which is
CC	the data collected with respect to the anti-neoplastic agent as a result
CC	of M1, and the data is sufficient to convey the chemical structure and/or
CC	properties of the agent. M1 can be used in the treatment of cancer such
CC	as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC	prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC	cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC	cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
CC	tumour
XX	
SQ	Sequence 1399 BP; 464 A; 273 C; 298 G; 364 T; 0 U; 0 Other;
	Query Match 91.5%; Score 1387.8; DB 6; Length 1399;
	Best Local Similarity 99.8%; Pred. No. 0;
	Matches 1389; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY	125 AGTGTGAATCTTCAGAGAAGAAATTCCTTTAGTCTTCGCAAGAGTGACAGATAAG 184
Db	1 AGTGTGAATCTTCAGAGAAGAAATTCCTTTAGTCTTCGCAAGAGTGACAGATAAG 60
QY	185 ACATCTTTCAAATGGCATGGTATCAGAAATTCCTCAGCAGGCCTCGTTATTGAAA 244
Db	61 ACATCTTTCAAATGGCAATGGTATCAGAAATTCCTCAGCAGGCCTCGTTATTGAAA 120
QY	245 ATGAAGAGCAGGAATATGTTCAAACCTGTGAAGTCATCCAAGGTGTCGGATCAGCG 304
Db	121 ATGAAGAGCAGGAATATGTTCAAACCTGTGAAGTCATCCAAGGTGTCGGATCAGCG 180
QY	305 TGAGCCCCCTATCCTACTTCATCCATCCTCGGATGCGTCTGCCATTAAGGCCATAA 364
Db	181 TGAGCCCCCTATCCTACTTCATCCATCCTCGGATGCGTCTGCCATTAAGGCCATAA 240
QY	365 TGGTTAAAGGTGTGGATGAAGCAACCATCTTAGCATTTCTAACCTAAGCGAAACAATGCAC 424
Db	241 TGGTTAAAGGTGTGGATGAAGCAACCATCTTAGCATTTCTAACCTAAGCGAAACAATGCAC 300
QY	425 AGCGTCAACMGATCAAGCAGCATATCTCCAGGAAACAGGAAAGCCCCTGGATGAAACAC 484
Db	301 AGCGTCAACAGATCAAGCAGCATATCTCCAGGAAACAGGAAAGCCCCTGGATGAAACAC 360
QY	485 TGAAGAAGGCCCTTACAGCTCACCTTGAGGAGGTTTGTTTAGCTCTGCTAAAAAATCCGAG 544
Db	361 TTGAAGAAGGCCCTTACAGCTCACCTTGAGGAGGTTTGTTTAGCTCTGCTAAAAAATCCGAG 420
QY	545 CGCAATTTGATGATGAACCTTCGTGTCCTGATGAAGGCCCTTGGAACTGATGAAGATA 604
Db	421 CGCAATTTGATGATGAACCTTCGTGTCCTGATGAAGGCCCTTGGAACTGATGAAGATA 480

ABX84511
ID ABX84511 standard; cDNA; 1399 BP.

XX AC ABX84511;

XX DT 14-AUG-2002 (first entry)

XX DE Human cDNA differentially expressed in granulocytic cells #1082.

XX KW Human; ss; granulocytic cell; DNA chip; bacterial infection;
viral infection; parasitic infection; protozoal infection;
fungal infection; sterile inflammatory disease; psoriasis;
rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
cardiac reperfusion injury; renal reperfusion injury; ARDS;
adult respiratory distress syndrome; inflammatory bowel disease;
Crohn's disease; ulcerative colitis; periodontal disease;
granulocyte activation; chronic inflammation; allergy.

XX OS Homo sapiens.

XX PN WO200228999-A2.

XX PD 11-APR-2002.

XX PF 03-OCT-2001; 2001WO-US030821.

XX PR 03-OCT-2000; 2000US-0237189P.

XX PA (GENE-) GENE LOGIC INC.

XX PI Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;

XX DR WPI; 2002-43528/46.

XX PT Detecting granulocyte activation by detecting differential expression of
PT genes associated with granulocyte activation, which serves as diagnostic
PT markers that is useful for monitoring disease states and drug toxicity.

XX PS Claim 1; SEQ ID NO 1082; 114pp; English.

CC The invention relates to detecting (M1) granulocyte (GC) activation
CC (GCA) by detecting the level of expression of gene(s) (Gs) identified by
CC DNA chip analysis as given in the specification, and comparing the
CC expression level to an expression level in an unactivated GC, where
CC differential expression of Gs is indicative of GCA. Also included are
CC modulating (M2) GA by contacting GC with an agent that alters the
CC expression of at least one gene in Gs; (2) screening (M3) for an agent
CC capable of modulating GCA or an inflammation (especially chronic) in a
CC tissue, an allergic response in a subject, exposure of a subject to a
CC pathogen or sterile inflammatory disease using the gene expression
CC profile; (3) detecting (M4) an inflammation (especially chronic) in a
CC tissue, an allergic response in a subject, exposure of a subject to a
CC pathogen or sterile inflammatory disease, by detecting the level of
CC expression in a sample of the tissue of gene(s) from Gs, where the level
CC of expression of the gene is indicative of inflammation; (4) treating
CC (M5) an inflammation (especially chronic) or in a tissue, an allergic
CC response in a subject, exposure of a subject to a pathogen or sterile
CC inflammatory disease, by contacting a tissue having inflammation with an
CC agent that modulates the expression of gene(s) from Gs in the tissue. M1
CC is useful for detecting GCA; M2 is useful for modulating GCA; M3 is useful
CC for screening an agent capable of modulating GCA preferably in an
CC inflammation in a tissue; M4 is useful for detecting an inflammation
CC (especially chronic) in a tissue, an allergic response in a subject,
CC exposure of a subject to a pathogen or sterile inflammatory disease (e.g.
CC psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis,
CC cardiac reperfusion injury, renal reperfusion injury, ARDS, adult
CC respiratory distress syndrome, inflammatory bowel disease, Crohn's
CC disease, ulcerative colitis, periodontal disease; also bacterial
CC infection, viral infection, parasitic infection, protozoal infection,
CC fungal infection and M5 is useful for treating one of the above
CC conditions. The present sequence represents a gene differentially
CC expressed in granulocytes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in

CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 1399 BP; 464 A; 273 C; 298 G; 364 T; 0 U; 0 Other;

Query Match 91.5%; Score 1387.8; DB 6; Length 1399;
Best Local Similarity 99.8%; Pred. NO. 0;
Matches 1389; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 125 AGTGTGAATCTTCAGAGAGAAATTTCTCTTTAGTTCTTTTGAAGAAGCTAGAGATAAG 184

DB 1 AGTGTGAATCTTCAGAGAGAAATTTCTCTTTAGTTCTTTTGAAGAAGCTAGAGATAAG 60

QY 185 ACATCTTTTCAAAAATGGCAATGGTATCAGATTCTCTCAAGCAGGCTGTTTATTGAA 244

DB 61 ACATCTTTTCAAAAATGGCAATGGTATCAGATTCTCTCAAGCAGGCTGTTTATTGAA 120

QY 245 ATGAAGAGCAGGAATATGTTTCAAACTGTGAAGTCATCCAAAGTGTTCCCGGATCAGCG 304

DB 121 ATGAAGAGCAGGAATATGTTTCAAACTGTGAAGTCATCCAAAGTGTTCCCGGATCAGCG 180

QY 305 TGAGCCCTCTATCTTCAATCCATCTCGGATGTCCTGCTTGCATTAAGGCCATAA 364

DB 181 TGAGCCCTCTATCTTCAATCCATCTCGGATGTCCTGCTTGCATTAAGGCCATAA 240

QY 365 TGGTTAAAGGTGTGATGAAGCAACCATCATTTGACATTTAACTAAGCGAAACAATGAC 424

DB 241 TGGTTAAAGGTGTGATGAAGCAACCATCATTTGACATTTAACTAAGCGAAACAATGAC 300

QY 425 AGGCTCAACAGATCAAAAGCAGCATATCTCCAGAAACAGAAAGCCCTGGATGAACAC 484

DB 301 AGGCTCAACAGATCAAAAGCAGCATATCTCCAGAAACAGAAAGCCCTGGATGAACAC 360

QY 485 TGAAGAAGCCCTTACAGTCACTTGAAGAGGTTGTTTACCTCTGCTTAAACATCCAG 544

DB 361 TTAAGAAGCCCTTACAGTCACTTGAAGAGGTTGTTTACCTCTGCTTAAACATCCAG 420

QY 545 CGCAATTTGATGCTGATGAATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 604

DB 421 CGCAATTTGATGCTGATGAATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480

QY 605 CTCTAATTTGATTTGGCATCAAGAACTTAAACAAGAAATCAGACATTAACAGGTTCT 664

DB 481 CTCTAATTTGATTTGGCATCAAGAACTTAAACAAGAAATCAGACATTAACAGGTTCT 540

QY 665 ACAGAGAGAACTTGAAGAGATCTGCCCAAGACATAACCTCAGACACATCTGGAGATT 724

DB 541 ACAGAGAGAACTTGAAGAGATCTGCCCAAGACATAACCTCAGACACATCTGGAGATT 600

QY 725 TTGGAAACGTTTGTCTTCTTTGCTAAGGGTGACCGATCTGAGGACTTTGGTGTGAATG 784

DB 601 TTGGAAACGTTTGTCTTCTTTGCTAAGGGTGACCGATCTGAGGACTTTGGTGTGAATG 660

QY 785 AAGACTTGGCTGATTTCAGATGCGAGGCTTGTATGAAGCAGGAGAAAGGAGAAAGGGA 844

DB 561 AAGACTTGGCTGATTTCAGATGCGAGGCTTGTATGAAGCAGGAGAAAGGAGAAAGGGA 720

QY 845 CAGACGTAACGTTGTTCAATACATCTTACACAGAGAGCTATCCACACTTCCACAG 904

DB 721 CAGACGTAACGTTGTTCAATACATCTTACACAGAGAGCTATCCACACTTCCACAG 780

QY 905 TGTTTCAGAAATACACCAAGTACAGTGAAGCATGAACAAAGTTTGGACCTTGGAGT 964

DB 781 TGTTTCAGAAATACACCAAGTACAGTGAAGCATGAACAAAGTTTGGACCTTGGAGT 840

QY 965 TGAAGGTGACATTGAGAATGCTTCAGAGCTATCGTGAAGTGGCCCAAGCAACACAG 1024

DB 841 TGAAGGTGACATTGAGAATGCTTCAGAGCTATCGTGAAGTGGCCCAAGCAACACAG 900

QY 1025 CTCTTCTTTCAGAGAGAGCTTTCATCAAGCCATCAAAAGGTTTGGAACTCGCCCAAGGAT 1084

DB 901 CTCTTCTTTCAGAGAGAGCTTTCATCAAGCCATCAAAAGGTTTGGAACTCGCCCAAGGAT 960

1085	YGATCAGGATTATGGTTTCCCGTTCTCGAATTCGACATGAATGATATCAAGCAATTCCTATC	1144
961	YGATCAGGATTATGGTTTCCCGTTCTCGAATTCGACATGAATGATATCAAGCAATTCCTATC	1020
1145	AGAAGATGATGCTACTCCCTTTGGCCAGCCCATCTGGATGAAACCAAGGAGAGTATG	1204
1021	AGAAAGATGATGCTACTCCCTTTGGCCAGCCCATCTGGATGAAACCAAGGAGATATG	1080
1205	AGAAAAATCCTGGTGGCTCTTTTGGAGGAAACTAAACATTCCTTCGATGGTCTCAAGCTA	1264
1081	AGAAAAATCCTGGTGGCTCTTTTGGAGGAAACTAAACATTCCTTCGATGGTCTCAAGCTA	1140
1265	TGATCAGAGAGACCTTTAAATATATATTTTCATCCTATAAGCTTAAATAGGAAAGTTTCTTC	1324
1141	TGATCAGAGAGACCTTTAAATATATATTTTCATCCTATAAGCTTAAATAGGAAAGTTTCTTC	1200
1325	AACAGGATTACAGTGTAGCTACCTACATGCTGAAAAATATAGCCTTTAAATCAATTTTTAT	1384
1201	AACAGGATTACAGTGTAGCTACCTACATGCTGAAAAATATAGCCTTTAAATCAATTTTTAT	1260
1385	ATTATAACTCTGTATATAGAGATAGTCCATTTTAAAAATGTTTTCCCAAAACCATTA	1444
1261	ATTATAACTCTGTATATAGAGATAGTCCATTTTAAAAATGTTTTCCCAAAACCATTA	1320
1445	AAACCCCTATACAAAGTTGTTCTTAGTAACAATACATGAGAAAGATGCTCTATGTAGCTGAAA	1504
1321	AAACCCCTATACAAAGTTGTTCTTAGTAACAATACATGAGAAAGATGCTCTATGTAGCTGAAA	1380
1505	TAAAAATGNCCTC	1516
1381	TAAAAATGACGTC	1392

RESULT 11

ABV78074

ID ID ABV78074 standard; DNA; 1399 BP.

XX AC ABV78074;

XX AC

DT 12-NOV-2002 (first entry)

XX DE

XX Hypoxia-regulated protein coding sequence #94.

XX DE

XX Cytostatic; vasotropic; tranquiliser; antiatherosclerotic; gene therapy;

XX KW antiinflammatory; vulnery; gynecological; ophthalmological; vaccine;

XX KW hypoxia; tumorigenesis; angiogenesis; apoptosis; cancer;

XX KW ischemic condition; reperfusion injury; retinopathy; neonatal stress;

XX KW preeclampsia; atherosclerosis; inflammatory condition; wound healing;

XX KW inflammation; erythropoiesis; hair loss; human; gene; ds.

XX OS Homo sapiens.

XX OS WO200246465-A2.

XX PN

XX 13-JUN-2002.

XX PD

XX 10-DEC-2001; 2001WO-GB005458.

XX XX

XX 08-DEC-2000; 2000GB-00030076.

XX PR 08-FEB-2001; 2001GB-00003156.

XX PR 25-OCT-2001; 2001GB-00025666.

XX XX

XX (OXFO-) OXFORD BIOMEDICA UK LTD.

XX FA

XX White JN, Mundy CR, Ward NR, Krige D, Kingsman SM, Harris RA;

XX PI Rayner W;

XX PI

XX WPI; 2002-627238/67.

XX DR

XX Identifying a gene involved in disease for treating hypoxia-regulated

XX PT conditions, comprises comparing the transcriptome/procome of two cell

XX PT types under different conditions and identifying a differentially

XX PT regulated gene.

327	QY	TCCATCTCGGATGTGGCTGCCCTTGCATAGGCCCAATAATGGTTAAAGGTGTGGATGAAGC	386
181	Db	TCCATCTCGGATGTGGCTGCCCTTGCATPAAGGCCAATAATGGTTAAAGGTGTGGATGAAGC	240
387	QY	AACCATATTGACATTTCTAACTTAAGCGAAACAAATGCGACAGCGTCAACAGATCAAAAGCAGC	446
241	Db	AACCATATTGACATTTCTAACTTAAGCGAAACAAATGCGACAGCGTCAACAGATCAAAAGCAGC	300
447	QY	ATATCTCCAGGAAACAGGAAAGCCCTCGATGAAACACTGAAAGAAAGCCCTTACAGGTCA	506
301	Db	ATATCTCCAGGAAACAGGAAAGCCCTCGATGAAACACTGAAAGAAAGCCCTTACAGGTCA	360
507	QY	CTTTGAGGAGTTGTTTTTGTAGTCTGTCTAAAACTCCAGCGCAATTTGATGCTGATGAAT	566
361	Db	CTTTGAGGAGTTGTTTTTGTAGTCTGTCTAAAACTCCAGCGCAATTTGATGCTGATGAAT	420
567	QY	TGCTGTGCCATCAAGGCGCTTGGAACTCATGAAGATACTCTAAATTGAGATTTTGGCATC	626
421	Db	TGCTGTGCCATCAAGGCGCTTGGAACTCATGAAGATACTCTAAATTGAGATTTTGGCATC	480
627	QY	AAGAACTTAAACAAAGAAATCAGAGACATTTAACGGGTCTACAGAGGAACTGAAAGAGAGA	686
481	Db	AAGAACTTAAACAAAGAAATCAGAGACATTTAACGGGTCTACAGAGGAACTGAAAGAGAGA	540
687	QY	TCTGGCCAAAGACATTAACCTCAGACATCTGAGAGATTTTCGAAACGCTTTCCTTCTCT	746
541	Db	TCTGGCCAAAGACATTAACCTCAGACATCTGAGAGATTTTCGAAACGCTTTCCTTCTCT	600
747	QY	TGCTAAGGGTGACCGATCTCAGGACCTTGGTGTGAATGAAGACTTGGCTGTATCAGATGC	806
601	Db	TGCTAAGGGTGACCGATCTCAGGACCTTGGTGTGAATGAAGACTTGGCTGTATCAGATGC	660
807	QY	CAGGGCCTTCTATCAGCAGGAGAAAGGAGAAAGGGGACAGACGTAAACCGTGTCAATAC	866
661	Db	CAGGGCCTTCTATCAGCAGGAGAAAGGAGAAAGGGGACAGACGTAAACCGTGTCAATAC	720
867	QY	CATCCTTTACCACCAGAAAGCTATCCCAACTTCCAGAGTGTTCAGAAATACACCAAGTA	926
721	Db	CATCCTTTACCACCAGAAAGCTATCCCAACTTCCAGAGTGTTCAGAAATACACCAAGTA	780
927	QY	CAGTAAGCATGACATGAACAAAGTTCTGGACCTGGAGTTGAAGGTGACATTGAGAAATG	986
781	Db	CAGTAAGCATGACATGAACAAAGTTCTGGACCTGGAGTTGAAGGTGACATTGAGAAATG	840
987	QY	CCTCACAGCTATCGTGAAGTGCGCCACAAGCAAAACAGACCTTCTTTTCGAGAGAGCTTCA	1046
841	Db	CCTCACAGCTATCGTGAAGTGCGCCACAAGCAAAACAGACCTTCTTTTCGAGAGAGCTTCA	900
1047	QY	TCAGGCCAATGAAGGTGTGAACTCCGCCATPAAGGCATTTGATCAGGATTAATGGTTTCCCG	1106
901	Db	TCAGGCCAATGAAGGTGTGAACTCCGCCATPAAGGCATTTGATCAGGATTAATGGTTTCCCG	960
1107	QY	TTCTGAAATTCACATGAATGATATCAAAAGATTCATCAGAAAGATGTATGGTATCTCCCT	1166
961	Db	TTCTGAAATTCACATGAATGATATCAAAAGATTCATCAGAAAGATGTATGGTATCTCCCT	1020
1167	QY	TTGCCAAGCCATCTCTGGATGAAACAAAGGAGAGATGAGAAATCTCTGGTGCCTTTTG	1226
1021	Db	TTGCCAAGCCATCTCTGGATGAAACAAAGGAGAGATGAGAAATCTCTGGTGCCTTTTG	1080
1227	QY	TGGAGGAACTTAAACATTTCCCTTGTATGTCTCAAGCTATGATCAGAAAGATTTAAATATA	1286
1081	Db	TGGAGGAACTTAAACATTTCCCTTGTATGTCTCAAGCTATGATCAGAAAGATTTAAATATA	1140
1287	QY	TATTTTTCATCTTAAAGCTTAAATAGGAAAGTTTCTCAACAGGATTAAGTGTAGCTAC	1346
1141	Db	TATTTTTCATCTTAAAGCTTAAATAGGAAAGTTTCTTCAACAGGATTAAGTGTAGCTAC	1200
1347	QY	CTACATGCTGAAAAATATAGCCTTTAAATCATTTTATATTAATTAACCTCTGTATAATAGAG	1406
1201	Db	CTACATGCTGAAAAATATAGCCTTTAAATCATTTTATATTAATTAACCTCTGTATAATAGAG	1260
1407	QY	ATAAGTCCATTTTTTAAAAATGTTTTTCCCCAAACCAATAAAACCCCTATACAAAGTTGTCTA	1466

1261 ATAAGTCCATTTTAAAAATGTTTCCCAACCCATAAAACCCATATACAAGTGTCTTA 1320
1467 GTAACAATACATGAGAAAGATGCTATGTAGTGTGAAATAAAATGNCGTC 1516
1321 GTAACAATACATGAGAAAGATGCTATGTAGTGTGAAATAAAATGACGTC 1370

RESULT 14
AAN60555
ID AAN60555 standard; cDNA; 1376 BP.
AC AAN60555;
XX
AC AAN60555;
XX
DT 08-AUG-1991 (first entry)
XX
DE Sequence of the cDNA insert of lambda-LC which codes for human
XX lipocortin.
DE
XX Anti-inflammatory agent; steroid mediation; arthritis therapy; allergy;
XX ds.
XX Homo sapiens.
XX
XX Location/Qualifiers
FH CDS 1..1092
FT /*tag= a
FT
XX
XX W08604094-A.
XX
XX 17-JUL-1986.
XX
XX 10-JAN-1986; 86WO-US000027.
XX
XX 10-JAN-1985; 85US-00690146.
XX 15-MAR-1985; 85US-00712376.
XX 14-AUG-1985; 85US-00765877.
XX 05-SEP-1985; 85US-00772892.
XX
XX (BIOJ) BIOGEN NV.
XX
XX Wallner BP, Pepinsky BR, Garwin JL, Schindler DG, Huang KS;
XX
XX WPI; 1986-196889/30.
XX P-FSDE; AAP60657.
XX
XX New human lipocortin-like polypeptide(s) - are obtd. by recombinant DNA
XX techniques and are antiinflammatory agents without usual side effects.
XX
XX Claim 2; Page 74-75 and Fig 4; 113pp; English.
XX
XX A human cDNA library of Escherichia coli cells contg. human macrophage
XX cDNA sequences inserted into a phage cloning vector was screened using
XX antisense oligonucleotide DNA probes corresp. to those regions of rat
XX phospholipase A2 inhibitor protein having minimal nucleotide degeneracy
XX (AAN60561-N60564) to obtain a sequence coding for human lipocortin (LC)
XX like polypeptide
XX
XX Sequence 1376 BP; 454 A; 271 C; 292 G; 359 T; 0 U; 0 Other;
XX

Query Match 90.0%; Score 1364.8; DB 1; Length 1376;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1366; Conservative 0; Mismatches 3; Indels 0; Gaps 0

148 TTTCTCTTTAGTCTTTCCAGAGAGGTAGATAAAGACACTTTTCAAAAATGGCAATG 207
1 TTTCTCTTTAGTCTTTCCAGAGAGGTAGATAAAGACACTTTTCAAAAATGGCAATG 60

208 GTATCAGAAATTCCTCAAGCAGCCTGGTTTATTGAAAAATGAAGACAGGAATATCTTCAA 267
61 GTATCAGAAATTCCTCAAGCAGCCTGGTTTATTGAAATGAAGACAGGAATATGTTCAA 120

268 ACTGTGAAGTCAATCAAAAGGTGGTCCCGGATCAGCGGTGAGCCCTATCTACCTTCAT 327

Db 121 ACTGTGAAGTCAATCAAAAGTGGTCCCGATCAGCGTGAGCCCTATCTACTCTCAAT 190
 Qy 328 CCATCTCTGGATGTCGCTTGCATTAAGGCCCAATATGTTAAAGGTGGAATGAAGCA 387
 Db 181 CCATCTCTGGATGTCGCTTGCATTAAGGCCCAATATGTTAAAGGTGGAATGAAGCA 240
 Qy 388 ACCATCATTTGACATTTCTAACTTAAGCGGAAACAAATSCACAGCGTCAACAGATCAAAAGCAGCA 447
 Db 241 ACCATCATTTGACATTTCTAACTTAAGCGGAAACAAATSCACAGCGTCAACAGATCAAAAGCAGCA 300
 Qy 448 TATCTCCAGGAAACAGGAAAGCCCTCGATGAATGAACACATGAAGAAAGCCCTTACAGTCAAC 507
 Db 301 TATCTCCAGGAAACAGGAAAGCCCTCGATGAATGAACACATGAAGAAAGCCCTTACAGTCAAC 360
 Qy 508 CTTGAGGAGTGTGTTTGTAGCTCTGCTTAAACTCTCAGCGCAATTTGATGCTGATGAATCT 567
 Db 361 CTTGAGGAGTGTGTTTGTAGCTCTGCTTAAACTCTCAGCGCAATTTGATGCTGATGAATCT 420
 Qy 568 CGTCTGCTCATGAAGGCGCTTGGAACTGATGAAGATCTCTAAATTGAGATTTGGGCATCA 627
 Db 421 CGTCTGCTCATGAAGGCGCTTGGAACTGATGAAGATCTCTAAATTGAGATTTGGGCATCA 480
 Qy 628 AGAATCAACAGAAATCAGAGACATTAACAGGCTCTACAGAGAGGAACTGAAGAGAGAT 687
 Db 481 AGAATCAACAGAAATCAGAGACATTAACAGGCTCTACAGAGAGGAACTGAAGAGAGAT 540
 Qy 688 CTGGCCAAAGACATTAACCTCAGACACATCTCGAGATTTTCGGAACGCTTTGCTTTCTCT 747
 Db 541 CTGGCCAAAGACATTAACCTCAGACACATCTCGAGATTTTCGGAACGCTTTGCTTTCTCT 600
 Qy 748 GCTAAGGTGACCGATCTGAGGACCTTGGTGTGATGATGAGATCTGGCTGATTCAGATGCC 807
 Db 601 GCTAAGGTGACCGATCTGAGGACCTTGGTGTGATGATGAGATCTGGCTGATTCAGATGCC 660
 Qy 808 AGGCGCTTGTATGAAGCAGGAGAAAGGAGAAAGGAGCAGACGCTAAAGCTGTTCATATACC 867
 Db 661 AGGCGCTTGTATGAAGCAGGAGAAAGGAGAAAGGAGCAGACGCTAAAGCTGTTCATATACC 720
 Qy 868 ATCTTACACACAGAGATATCCACAACTTCGAGAGTGTTCAGAAATACACCAAGTAC 927
 Db 721 ATCTTACACACAGAGATATCCACAACTTCGAGAGTGTTCAGAAATACACCAAGTAC 780
 Qy 928 AGTAGCATGACATCAACAAAGTCTCGACCTGGAGTTGAAAGGTGACATGAGAAATGC 987
 Db 781 AGTAGCATGACATCAACAAAGTCTCGACCTGGAGTTGAAAGGTGACATGAGAAATGC 840
 Qy 988 CTCACAGCTATCGTGAAGTGGCCCAACAGCAACAGCTTTCTTTGACAGAAAGCTTCAT 1047
 Db 841 CTCACAGCTATCGTGAAGTGGCCCAACAGCAACAGCTTTCTTTGACAGAAAGCTTCAT 900
 Qy 1048 CAAGCCATGAAGGTTGGTAATCGGACTCGCATAGGCAATGATCAGGATATGTTTCCCGT 1107
 Db 901 CAAGCCATGAAGGTTGGTAATCGGACTCGCATAGGCAATGATCAGGATATGTTTCCCGT 960
 Qy 1108 TCTGAAATTCACATGAATGATCAAGCAATTCATCAGAAAGATGATGATGATCTCCCTT 1167
 Db 961 TCTGAAATTCACATGAATGATCAAGCAATTCATCAGAAAGATGATGATGATCTCCCTT 1020
 Qy 1168 TGCAGAGCATCTCGATGAATCAACAGAGAGATGATGAGAAATCCTGCTGCTTTGT 1227
 Db 1021 TGCAGAGCATCTCGATGAATCAACAGAGAGATGATGAGAAATCCTGCTGCTTTGT 1080
 Qy 1228 GGAGGAAACTAAACATTCCTCTGATGCTCAAGCTATGATCAGAAAGCTTTAAATATAT 1287
 Db 1081 GGAGGAAACTAAACATTCCTCTGATGCTCAAGCTATGATCAGAAAGCTTTAAATATAT 1140
 Qy 1288 ATTTTCATCTTAAGCTTAATAGAAAGTTCTTCAACAGATTAAGTGTACTACC 1347
 Db 1141 ATTTTCATCTTAAGCTTAATAGAAAGTTCTTCAACAGATTAAGTGTACTACC 1200
 Qy 1348 TACATCTGAAAATATAGCTTTAAATCAATTTTATATATATACTCTGTAATATAGAGA 1407

Db 1201 TACATCTGAAAATATAGCCCTTAAATCAATTTTATATATTAATCTGTATATAAGAGA 1260
 Qy 1408 TAAGTCCATTTTAAAAATGTTTTCCCAACCAATAAAACCTATACAAGTTGTTCTAG 1467
 Db 1261 TAAGTCCATTTTAAAAATGTTTTCCCAACCAATAAAACCTATACAAGTTGTTCTAG 1320
 Qy 1468 TAACAATACATGAGAAAGATGCTATGATGCTGAGAAATAAAAATGCTC 1516
 Db 1321 TAACAATACATGAGAAAGATGCTATGATGCTGAGAAATAAAAATGACGTC 1369

RESULT 15
 AAQ23216
 ID AAQ23216 standard; cDNA; 1376 BP.
 XX
 AC AAQ23216;
 XX
 DT 25-MAR-2003 (revised)
 DT 13-MAY-1992 (first entry)
 XX
 Lambda LC insert encoding human lipocortin.
 XX
 KW Antiinflammatory; arthritis; allergy; dermatology;
 KW phospholipase A2 inhibitor; ss.
 XX
 OS Homo sapiens.
 XX
 Key Location/Qualifiers
 CDS 1..1092
 FT /*tag= a
 FT sig_peptide 1..57
 FT /*tag= b
 FT mat_peptide 58..1089
 FT /*tag= c
 FT /*product= "lipocortin"
 FT misc_feature 81..111
 FT /*tag= d
 FT /*label= lipo 16
 FT /*note= "probe"
 XX
 PN US5081019-A.
 XX
 PD 14-JAN-1992.
 XX
 PF 02-MAY-1990; 90US-00519256.
 XX
 PR 10-JAN-1985; 85US-00690146.
 PR 15-MAR-1985; 85US-00712376.
 PR 14-AUG-1985; 85US-00765877.
 PR 05-SEP-1985; 85US-00772892.
 PR 06-MAR-1986; 86US-00837019.
 PR 22-FEB-1989; 89US-00314316.
 XX
 PA (BIOJ) BIOGEN INC.
 XX
 PI Wallner BP, Pepinsky RB;
 XX
 DR WPI; 1992-048295/06.
 DR P-PSDB; AAR22402.
 XX
 PT Deoxyribonucleic acid encoding lipo corticoid polypeptide(s) - for
 PT producing antiinflammatory agents for treating arthritic, allergic,
 PT dermatologic ophthalmic and collagen diseases.
 XX
 PS Disclosure; Fig 4; 56pp; English.
 CC
 CC The sequence was obtd. from 14 overlapping clones isolated from a cDNA
 CC library prepd. from polyA+ mRNA from human macrophage cell line U937. The
 CC library in lambda gt10, was initially screened using probes (see
 CC AAQ2319-22) designed from the corresponding rat sequence. Twenty
 CC positive plaques were rescreened and two clones were isolated. lambda9-
 CC 111 and lambda04-211. Restriction fragments from the clones were subcloned
 CC to prepare 3 plasmids which were sequenced. A 480 bp fragment of one of

CC these plasmids, pL5/20, was used as probe to rescreen the original
CC library. Seventy-two plaques were isolated and analysed by Southern
CC blotting using a 30mer (lipo 16; tag d) as a probe. The 14 clones
CC isolated were sequenced and used to determine the full sequence encoding
CC lipocortin. Recombinant lipocortin prep. using the DNA sequence has
CC antiinflammatory activity and can be used for the treatment of arthritic,
CC allergic, dermatologic, opthalmic and collagen diseases. See also
CC AAQ23217-40 and AAQ23294. (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ

Sequence 1376 BP; 454 A; 271 C; 292 G; 359 T; 0 U; 0 Other;

Query Match	90.08;	Score 1364.8;	DB 2;	Length 1376;
Best Local Similarity	99.88;	Pred. No. 0;		
Matches 1366;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;
148	TTTCTCTTTAGTTCTTTGCAAGAGGTAGAGATAAAGACACTTTTTCAAAAATGGCAATG	207		
1	TTTCTCTTTAGTTCTTTGCAAGAGGTAGAGATAAAGACACTTTTTCAAAAATGGCAATG	60		
208	GTATCAGAAATCCTCRAGCAGGCTGTTTATTGAAATGAAGCAGGAGGAATATGTTCAA	267		
61	GTATCAGAAATCCTCRAGCAGGCTGTTTATTGAAATGAAGCAGGAGGAATATGTTCAA	120		
268	ACTGTGAAGTCATCCAAAGGTGTTCCGGATCAGCGGTAGCGCCCTATCCTACCTTCAAT	327		
121	ACTGTGAAGTCATCCAAAGGTGTTCCGGATCAGCGGTAGCGCCCTATCCTACCTTCAAT	180		
328	CCATCTCGGATGCTGCTGCTGCATTAAGCCATTAATGTTAAAGGTGCGATGAAGCA	387		
181	CCATCTCGGATGCTGCTGCTGCATTAAGCCATTAATGTTAAAGGTGCGATGAAGCA	240		
388	ACCATCATTTGACATTTCTAACTTAAGCGAAACAATGCACAGCGTCAACAGATCAAAGCAGCA	447		
241	ACCATCATTTGACATTTCTAACTTAAGCGAAACAATGCACAGCGTCAACAGATCAAAGCAGCA	300		
448	TATCTCAGGAAACAGAAAGCCCTGGATGAACACTGAAGAGCCCTTACAGTCTAC	507		
301	TATCTCAGGAAACAGAAAGCCCTGGATGAACACTGAAGAGCCCTTACAGTCTAC	360		
508	CTTGAGGAGGTGTTGTTTAGCTCTGCTAAAAAATCCAGCGCAATTTGATGCTGATGAAC	567		
361	CTTGAGGAGGTGTTGTTTAGCTCTGCTAAAAAATCCAGCGCAATTTGATGCTGATGAAC	420		
568	CGTGTGCTATGAAGGCGCTTGGAACTGATGAAGATCTCTAATGAGATTTTGGCATCA	627		
421	CGTGTGCTATGAAGGCGCTTGGAACTGATGAAGATCTCTAATGAGATTTTGGCATCA	480		
628	AGAACTAAAGAAATCAGAGACATTAACAGGGTCTACAGAGAGGAACCTGAAGAGAGAT	687		
481	AGAACTAAAGAAATCAGAGACATTAACAGGGTCTACAGAGAGGAACCTGAAGAGAGAT	540		
688	CTGGCCAAAGACATAA CTTCTAGACACATCTGGAGATTTTCGGAAACGCTTTGCTTCTTT	747		
541	CTGGCCAAAGACATAA CTTCTAGACACATCTGGAGATTTTCGGAAACGCTTTGCTTCTTT	600		
748	GCTAAGGTGACCGATCTGAGGACTTTGGTGTGAATGAAGACTTTGGCTGATTCAGATGCC	807		
601	GCTAAGGTGACCGATCTGAGGACTTTGGTGTGAATGAAGACTTTGGCTGATTCAGATGCC	660		
808	AGGGCCTTTGATGAAGCAGGAGAAAGGAGAAAGGGGACAGACGTAAACGTTTCAATACC	867		
661	AGGGCCTTTGATGAAGCAGGAGAAAGGAGAAAGGGGACAGACGTAAACGTTTCAATACC	720		
868	ATCCTTACACAGAGACTTCCACACTTCGACAGAGTGTTCAGAAATACACCAAGTAC	927		
721	ATCCTTACACAGAGACTTCCACACTTCGACAGAGTGTTCAGAAATACACCAAGTAC	780		
928	AGTAAGCATGACATGAACAAAGTTCTGGACCTGGAGTTGAAAGGTGACATTTGAGAAATGC	987		
781	AGTAAGCATGACATGAACAAAGTTCTGGACCTGGAGTTGAAAGGTGACATTTGAGAAATGC	840		
988	CTCAGAGCTATCGTGAAGTGGCCACAGCAACACAGCTTTCTTTGAGAGAGCTTTCAT	1047		

Db	841	CTCACGCTATCGTGAAGTGGCCACCAAGCAAAACCAGCTTTCTTTGCAGAGAGCTTCAT	900
Qy	1048	CAAGCCATGAAGGTGTGGAACTCGCCATAAGGCATGATCAGGATATGTTTCCCGT	1107
Db	901	CAAGCCATGAAGGTGTGGAACTCGCCATAAGGCATGATCAGGATATGTTTCCCGT	960
Qy	1108	TCTGAAATTCACATGAATGATATCAAAAGCATTTCTATCAGAAGATGTATGTTATCTCCCTT	1167
Db	961	TCTGAAATTCACATGAATGATATCAAAAGCATTTCTATCAGAAGATGTATGTTATCTCCCTT	1020
Qy	1168	TGCCAAGCCATCTCGGATGAACCAAAAGGAGAGTATGAGAAAATCCTGGTGGCTCTTTGT	1227
Db	1021	TGCCAAGCCATCTCGGATGAACCAAAAGGAGAGTATGAGAAAATCCTGGTGGCTCTTTGT	1080
Qy	1228	GGAGGAACTAAACATTCCTTGATGCTCAAGCTATGATCAGAAGACTTTTAATTATAT	1287
Db	1081	GGAGGAACTAAACATTCCTTGATGCTCAAGCTATGATCAGAAGACTTTTAATTATAT	1140
Qy	1288	ATTTTCATCTTATTAAGCTTAAATAGGAAAGTTCTTCAACAGGATTAACAGTGTAGCTACC	1347
Db	1141	ATTTTCATCTTATTAAGCTTAAATAGGAAAGTTCTTCAACAGGATTAACAGTGTAGCTACC	1200
Qy	1348	TACATGCTGAAATATAGCCTTTAAATCATTTTATATTAATTAATTAATTAATTAATTAAT	1407
Db	1201	TACATGCTGAAATATAGCCTTTAAATCATTTTATTAATTAATTAATTAATTAATTAAT	1260
Qy	1408	TAAGTCCATTTTTTAAAAATGTTTTTCCCAAAACCATAAAAACCTTATCAAGTGTCTTAG	1467
Db	1261	TAAGTCCATTTTTTAAAAATGTTTTTCCCAAAACCATAAAAACCTTATCAAGTGTCTTAG	1320
Qy	1468	TAACTATACATGAGAAAGATGCTCTATGCTAGCTGAAATTAATTAATTAATTAATTAAT	1516
Db	1321	TAACTATACATGAGAAAGATGCTCTATGCTAGCTGAAATTAATTAATTAATTAATTAAT	1369

Search completed: March 10, 2004, 10:59:06
Job time : 373.965 secs

Db 346 TTTCCAGGAGCTGCGAGTTTGTGAGTGAACAGACGTGAAGGTCAACATCATGTGACGA 405
QY 3053 CCGCTGAGAGTGAAGTGAACCGGCTACCGTGTGATGTATCCCGCTCAACTGTGCG 3112
Db 406 CCGCTGAGAGTGAACCGGCTACCGTGTGATGTATCCCGCTCAACTGTGCG 465
QY 3113 GAGCAGGAGGAGGCGCCATGAGAGGAACCTTTGAGAGTCAACCGGCTGTCC 3172
Db 466 GAGCAGGAGGAGGCGCCATGAGAGGAACCTTTGAGAGTCAACCGGCTGTCC 525
QY 3173 CCGGAGGTCACTTATTAATTCAAGTCTTTGAGTGAAGCCATGAGAGGAGAGCAAGCCT 3232
Db 526 CCGGAGGTCACTTATTAATTCAAGTCTTTGAGTGAAGCCATGAGAGGAGAGCAAGCCT 585
QY 3233 CTGACTGCTCAAGAGAACCAACTGAGTGTCCCACTAACCTCCAG--TTTGTCAATGA 3291
Db 586 CTGACTGCTCAAGAGAACCAACTGAGTGTCCCACTAACCTCCAGTTTGTCAATGA 645
QY 3292 AACTGATTCTACTGTCTGTGAGATGAGTCACTCCAGGCGCCAGATTAACAGATACCG 3351
Db 646 AACTGATTCTACTGTCTGTGAGATGAGTCACTCCAGGCGCCAGATTAACAGATACCG 705
QY 3352 ACTGACCGTGGGCTTACCCGGAAGAGACAGCCGAGGAGTAAATGTGGTCCCTGT 3411
Db 706 ACTGACCGTGGGCTTACCCGGAAGAGGCGAGCCGAGGAGTAAATGTGGTCCCTGT 765
QY 3412 CTCAGATACCCACTGAGAGATCTGACGCTGATCTGAGTACACCGTATCCCTGTGCG 3471
Db 766 CTCAGATACCCACTGAGAGATCTGACGCTGATCTGAGTACACCGTATCCCTGTGCG 825
QY 3472 CATTAAGGGCAACCAAGAGAGCCCAAGCACTGAGTCTTTACCACTGACGCTGG 3531
Db 826 CATTAAGGGCAACCAAGAGAGCCCAAGCACTGAGTCTTTACCACTGACGCTGG 885
QY 3532 GAGCTTATTCCACTTACAAACCGAGGTGACTGAGACCACTTGTGATCAATGAGAC 3591
Db 886 GAGCTTATTCCACTTACAAACCGAGGTGACTGAGACCACTTGTGATCAATGAGAC 945
QY 3592 GCCT 3595
Db 946 GCCT 949

Search completed: March 7, 2004, 21:39:02
Job time : 16937 secs


```
/clone="CS0D1081V003"
/tissue type="PLACENTA COT 25-NORMALIZED"
/clone lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 11.4%; Score 888; DB 13; Length 954;
Best Local Similarity 97.5%; Pred. No. 4e-244;
Matches 930; Conservative 0; Mismatches 21; Indels 3; Gaps 3;

2Y 6334 CGTACCCACCCCTGGGTATGACACTGGAATGTATTACGCTTCCTGGCACTTCGTGCA 6393
2b 954 CGTACCCACCACTGGGTATGACACTGGAATGTATTACGCTTCCTGGCACTTCGTGCA 895
2Y 6394 GCRA-CCGAGTGTGGCAACAATGATCTTTG-AGGAACATGTTTATAGGCGGACCA 6451
2b 894 GCRAACCCAGTGTGGCGACAAATGATCTTNTGNAGNACATGTTTAGNTGGACCA 835
2Y 6452 -CGCCCCACAAACCGCCACCCCCATAGGCTATAGGCAAGCAACCATACCCGCGAATGAG 6510
2b 834 CCGCCCCACAAACCGGACGCCATAGGCTATAGGCAAGCAACCATACCCGCGAATGAG 775
2Y 6511 ACAGAGCTCTCTCTCAGACACCATCTCATGGGCCCATCCAGACACTTCTGAGTA 6570
2b 774 ACAGAGAGCTCTCTCTCAGACACCATCTCATGGGCCCATCCAGACACTTCTGAGTA 715
2Y 6571 CATCATTTTCATGTCATCTCTGTCAGTATGAGAACCCCTTACGTTTACGAGTTCCTCG 6630
2b 714 CATCATTTTCATGTCATCTCTGTCAGTATGAGAACCCCTTACGTTTACGAGTTCCTCG 655
2Y 6631 AACTTTCACAGTGCCTACTCTGACAGCCCTCAGAGAGTGCACCTACACATCATCT 6690
2b 654 AACTTTCACAGTGCCTACTCTGACAGCCCTCAGAGAGTGCACCTACACATCATCT 595
2Y 6691 GGAGGCACTGAAGACACAGAGGCAATAGGTTTGGGAAGAGTGTGTACCTGGGCAA 6750
2b 594 GGAGGCACTGAAGACACAGAGGCAATAGGTTTGGGAAGAGTGTGTACCTGGGCAA 535
2Y 6751 CTCTGTCAAGAGGCTTGAACCACTACGATGCTGTTGACCCCTTACACAGT 6810
2b 534 CTCTGTCAAGAGGCTTGAACCACTACGATGCTGTTGACCCCTTACACAGT 475
2Y 6811 GTCCCATTTATCCGTTGAGATGATGGGAACGAATGCTCTGAATCAGGCTTTAAACTGTT 6870
2b 474 TTCCCATTTATCCGTTGAGATGATGGGAACGAATGCTCTGAATCAGGCTTTAAACTGTT 415
2Y 6871 GTGCCAGTGTAGGCTTTGAAAGTGTCTATTTTCAGATGTATCTATGATGTTGCCA 6930
2b 414 GTGCCAGTGTAGGCTTTGAAAGTGTCTATTTTCAGATGTATCTATGATGTTGCCA 355
2Y 6931 TGCAATGGTGTGAACACTACAGATTGAGAGAGTGGGACCGTCAGGAGAAATGGCCA 6990
2b 354 TGCAATGGTGTGAACACTACAGATTGAGAGAGTGGGACCGTCAGGAGAAATGGCCA 295
2Y 6991 GATGATGAGTGCACATGCTCTTGGGAACGGAAGAGAAATTCAGTGTGACCCCTCATGA 7050
2b 294 GATGATGAGTGCACATGCTCTTGGGAACGGAAGAGAAATTCAGTGTGACCCCTCATGA 235
2Y 7051 GGCAACGTTTACGATGATGGGAGACATACCCAGTAGGAGACAGTGGCAGAGGAGATA 7110
2b 234 GGCAACGTTTACGATGATGGGAGACATACCCAGTAGGAGACAGTGGCAGAGGAGATA 175
2Y 7111 TCTCGGTGCCATTTGCTCCTGCACATGCTTTTGGAGGCGCAGCGGGCTTGGCGCTGTGACAA 7170
2b 174 TCTCGGTGCCATTTGCTCCTGCACATGCTTTTGGAGGCGCAGCGGGCTTGGCGCTGTGACAA 115
2Y 7171 CTGGCGAGACTGGGGGTGAACCCAGTCCCGAGAGGCACTACTGGCCAGTCCCTACACCA 7230
2b 114 CTGGCGAGACTGGGGGTGAACCCAGTCCCGAGAGGCACTACTGGCCAGTCCCTACACCA 55
```

```
7231 GTATTTTCAGAGATACCATCAGAGAACAAACACCTAATGTTTAATTGCCCAATTGA 7284
Db 54 GTATTTTCAGAGATACCATCAGAGAACAAACACCTAATGTTTAATTGCCCAATTGA 1

RESULT 15
AL544172 949 bp mRNA linear EST 31-MAY-2003
LOCUS AL544172 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA
DEFINITION clone CS0D1020YB20 5-PRIME, mRNA sequence.
ACCESSION AL544172
VERSION AL544172
KEYWORDS EST.
SOURCE AL544172.2 GI:31266017
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 949)
Li, W.B., Gruber, C., Jeese, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 15, 2001 this sequence version replaced gi:12876651.
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 1413.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0D1020DA10Q1&cluster=1413.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue Genoscope sequence ID : CS0D1020DA10Q1.
FEATURES
Location/Qualifiers
1..949
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1020YB20"
/tissue type="PLACENTA COT 25-NORMALIZED"
/clone lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 11.4%; Score 887.2; DB 9; Length 949;
Best Local Similarity 99.6%; Pred. No. 5.8e-244;
Matches 900; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

2Y 2693 ACAGCGCTGATGCCCCCTCTGACCCGACTGTGACCAAGTTGATGACCTCAATTGTT 2752
2b 46 ATAGCGCTGATGCCCCCTCTGACCCGACTGTGACCAAGTTGATGACCTCAATTGTT 105
2Y 2753 GTTCCTGGACGACGACCCCGCTCCCATACAGGTTACAGATAGTCTATTGCGCATCA 2812
2b 106 GTTCCTGGACGACGACCCCGCTCCCATACAGGTTACAGATAGTCTATTGCGCATCA 165
2Y 2813 GTAGAAGGTAGCAGACAGAACTCAACCTTCTGAAACTGCAAACTCCGTCACCTCAGT 2872
2b 166 GTAGAAGGTAGCAGACAGAACTCAACCTTCTGAAACTGCAAACTCCGTCACCTCAGT 225
2Y 2873 GACTTGCACCTGGTGTTCAGTATACATCATCTATCTGCTGGAGAGAAATCAAGAA 2932
2b 226 GACTTGCACCTGGTGTTCAGTATACATCATCTATCTGCTGGAGAGAAATCAAGAA 285
2Y 2933 AGTACACCTGTTGTCTATTCAACAGAAACCACTGGCACCCCAACGCTCAGATGAGTGC 2992
2b 286 AGTACACCTGTTGTCTATTCAACAGAAACCACTGGCACCCCAACGCTCAGATGAGTGC 345
2Y 2993 TCTCCAGGAGCTGCTGCTTTGTGGAGTGCAGACAGCTGAAGGTCACCATCATGTGGACA 3052
```

```

DEFINITION AL551753 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1062YH09 5-PRIME, mRNA sequence.
ACCESSION AL551753
VERSION AL551753.2 GI:31273569
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
JOURNAL Full-length cDNA libraries and normalization
COMMENT On Feb 15, 2001 this sequence version replaced gi:12890006.
Contact: Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 1413.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0D1062YH09&iscluster=1413.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0D1062YH09P1.
FEATURES
source
Location/Qualifiers
1..1039
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1062YH09"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (3r)
primer. Five prime end enriched, double-strand cDNA was
digested with NotI and cloned into the NotI and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 11.4%; Score 888.2; DB 9; Length 1039;
Best Local Similarity 94.3%; Pred. No. 3.8e-244;
Matches 931; Conservative 36; Mismatches 15; Indels 5; Gaps 5;
QY 6587 CTTGTTGGACATGATGAGAACCCCTACAGTTCAGGTTCTGMACTTCTACAGTGC 6646
DB 55 CCGGGATGCACTGATGAGAACCCCTACAGTTCAGGTTCTGMACTTCTACAGTGC 114
QY 6647 ACTCTGACAGGCTCACCAGAGGTGCCACCTACACATCATAGTGGAGGCACTGAAAGAC 6706
DB 115 ACTCTGACAGGCTCACCAGAGGTGCCACCTACACATCATAGTGGAGGCACTGAAAGAC 174
QY 6707 CAGCAGAGCATTAAGTTTCGGGAAGAGGTGTTACCGTGGGCACTCTGTCAAGAGGC 6766
DB 175 CAGCAGAGCATTAAGTTTCGGGAAGAGGTGTTACCGTGGGCACTCTGTCAAGAGGC 234
QY 6767 TTGAACCACTACGATGACCTGCTTGGACCTTACAGTTCACAGTTCACATATGCGCGT 6826
DB 235 TTGAMCCAMCCTACGATGACCTGCTTGGACCTTACAGTTCACATATGCGCGT 294
QY 6827 GGAGATGAGTGGGAACGAATGTCTGAATCAGCGCTTAACTGTTGTGTCAGTGTAGGC 6886
DB 295 GGAGATGAGTGGGAACGAATGTCTGAATCAGCGCTTAACTGTTGTGTCAGTGTAGGC 354
QY 6887 TTGGAAGTGGTCATTTTCAGATGTGATTCATCTAGTGTGTCATGACATGTTGTGAAC 6946
DB 355 TTGGAAGTGGTCATTTTCAGATGTGATTCATCTAGTGTGTCATGACATGTTGTGAAC 414
QY 6947 TACAAGATTGGAGAGAGTGGGACCGTCAGGAGAAAATGGCCAGATGATGAGTGCACA 7006
DB 415 TACAAGATTGGAGAGAGTGGGACCGTCAGGAGAAAATGGCCAGATGATGAGTGCACA 474
QY 7007 TGTCTTGGAAACGGAAGAGGAAATTCAGTGTGACCTCATGAGGCACTGTTACGAT 7066

```

```

DB 475 TGTCTTGGAAACGGAAGAGGAAATTCAGTGTGACCTCATGAGGCAACGTTATGAT 534
QY 7067 GATGGGAAGACATACACAGTAGGAAACAGTGGCAGAGGAATATCTCGTGCCATTGCG 7126
DB 535 GATGGGAAGACATACACAGTAGGAAACAGTGGCAGAGGAATATCTCGTGCCATTGCG 594
QY 7127 TCTTCACATGCTTTGGAGGCCAGCGGCTGTGCAACCTGCGGAGAGACTGGG 7186
DB 595 TCTTCACATGCTTTGGAGGCCAGCGGCTGTGCAACCTGCGGAGAGACTGGG 654
QY 7187 GGTGAACCCAGTCCGAAAGGCACTACTGGCCAGTCTCAACCAAGTATTTCTCAGAGATAC 7246
DB 655 GGTGAACCCAGTCCGAAAGGCACTACTGGCCAGTCTCAACCAAGTATTTCTCAGAGATAC 713
QY 7247 CATCAGAGAAACACACTAATGTTAATGTCCTCAATGAGTGTCTCATGCTTTAGATGTA 7306
DB 714 CATCAGAGAAACACACTAATGTTAATGTCCTCAATGAGTGTCTCATGCTTTAGATGTA 773
QY 7307 CAGGCTGACAGAGAGATTCGCGAGAGTAAATCATCTTTCCAAATCCAGAGGAACAGCAT 7366
DB 774 CAGGCTGACAGAGAGATTCGCGAGAGTAAATCATCTTTCCAAATCCAGAGGAACAGCAT 833
QY 7367 GTCTCTCGCAGAGTCCATCTAACTGGAGAGTGTAGCAGACCCAGCTTAGAGTTCT 7426
DB 834 GTCTCTCGCAGAGTCCATCTAACTGGAGAGTGTAGCAGACCCAGCTTAGAGTTCT 892
QY 7427 TCTTCTCTTTTAAGCCCTTCTCTGAGAGAGTGTCTCCAGCTTCAGCTCAACTCACAG 7486
DB 893 TCTTCTCTTTTAAGCCCTTCTCTGAGAGAGTGTCTCCAGCTTCAGCTCAACTCACAG 951
QY 7487 CTCTCTCAAGATCACCCTGGAGAGTGTCTGAGGTTTCTCATTAATGAGGCTGCACA 7546
DB 952 CTCTCTCAAGATCACCCTGGAGAGTGTCTGAGGTTTCTCATTAATGAGGCTGCACA 1010
QY 7547 TTGCTGTCTGCTCGAAGTATTCAG 7573
DB 1011 TTGCTGTCTGCTCGAAGTATTCAG 1035
RESULT 14
BX350259/c 954 bp mRNA linear EST 05-MAY-2003
LOCUS BX350259 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0D1081Y003 3-PRIME, mRNA sequence.
ACCESSION BX350259
VERSION BX350259.1 GI:30375450
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
JOURNAL Full-length cDNA libraries and normalization
COMMENT Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 1413.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0BA1029ZD10_CS02764_1&iscluster=1413.r.
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0BA1029ZD10_CS02764_1.
FEATURES
source
Location/Qualifiers
1..954
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

```


Best Local Similarity 98.3%; Pred. No. 1.4e-247; Matches 941; Conservative 0; Mismatches 13; Indels 3; Gaps 3;			
3548	TAACAACACCGAGTGACTGAGACCCACCAATTGTATCATATGGACGCTGTCTCAAGAATT 3607		
34	TACCGGTCCGAATTCCTCGGATACCAATTGTATCATATGGACGCTGTCTCAAGAATT 93		
3608	GGTTTAACTGGGTGTATCGACCAAGCCAGGAGAGGACACAC-GAGAAGTGACTTC 3666		
94	GGTTTAACTGGGTGTATCGACCAAGCCAGGAGAGGACACACTTGAAGTGTACTTC 153		
3667	AGACTCAGGAAGCATCGTTGTGTCGGCTTGACTCCAGGAGTAAATACGTCTACACCAT 3726		
154	AGACTCAGGAAGCATCGTTGTGTCGGCTTGACTCCAGGAGTAAATACGTCTACACCAT 213		
3727	CCAGTCTCAGAGATGGACAGAAAGAGATCGGCCAATTGTAAACAAAGTGTGCACCC 3786		
214	CCAGTCTCAGAGATGGACAGAAAGAGATCGGCCAATTGTAAACAAAGTGTGCACCC 273		
3787	ATTGTCTCCACCAACAACTTGCACTCGAGGCAAACTCCACACTCGAGTGTCCACAGT 3846		
274	ATTGTCTCCACCAACAACTTGCACTCGAGGCAAACTCCACACTCGAGTGTCCACAGT 333		
3847	CTCTGGGAGGAGGACCAACCCACAGACATTACTGGTTATAGAAATTACCAACCCCTAC 3906		
334	CTCTGGGAGGAGGACCAACCCACAGACATTACTGGTTATAGAAATTACCAACCCCTAC 393		
3907	AAACGGCCACGAGGAAATCTTTGGAGAGTGTGCATCTGCTGATCAGAGCTCTTCGAC 3966		
394	AAACGGCCACGAGGAAATCTTTGGAGAGTGTGCATCTGCTGATCAGAGCTCTTCGAC 453		
3967	TTTTGTAACTCAGTCCCGGCTGGAGTACAAATGTCTCAGTGTTTACACTGTCAAGATGA 4026		
454	TTTTGTAACTCAGTCCCGGCTGGAGTACAAATGTCTCAGTGTTTACACTGTCAAGATGA 513		
4027	CAAGAAAGTGTCCCTATCTGTATACCATCATCCAGCTGTTCCTCTCCACTGACCT 4086		
514	CAAGAAAGTGTCCCTATCTGTATACCATCATCCAGCTGTTCCTCTCCACTGACCT 573		
4087	GCGATTCACCAACATTGGTCCAGACACCATGGTGTCACTGGGCTCCACCCCATCCAT 4146		
574	GCGATTCACCAACATTGGTCCAGACACCATGGTGTCACTGGGCTCCACCCCATCCAT 633		
4147	TGATTTAACCAACTTCCTGGTGGTTATCTCCTCTGTGAAAAATGAGGAAGATGTTGAGA 4206		
634	TGATTTAACCAACTTCCTGGTGGTTATCTCCTCTGTGAAAAATGAGGAAGATGTTGAGA 693		
4207	GTTCGCAATTTCTCTCAGACAAATCGAGTGTCTTACAAATCTCTCTGGTGTGACAGA 4266		
694	GTTCGCAATTTCTCTCAGACAAATCGAGTGTCTTACAAATCTCTCTGGTGTGACAGA 753		
4267	ATATGTAGTAGTGTCTCCAGTGTCTACGAAACAAATCAGAGACACACTCTTAGAGGAG 4326		
754	ATATGTAGTAGTGTCTCCAGTGTCTACGAAACAAATCAGAGACACACTCTTAGAGGAG 813		
4327	ACAGAAACAGGCTTGATTTCCCAACCTGGCAATTGACTTTTCTGATATTACTGCCAATC 4386		
814	ACAGAAACAGGCTTGATTTCCCAACCTGGCAATTGACTTTTCTGATATTACTGCCAATC 873		
4387	TTTTTACTGTGCACTGGATTCTCCTCGAGCACCACTACTGGCTACAGGATCCGCCATCA 4446		
874	TTTTTACTGTGCACTGGATTCTCCTCGAGCACCACTACTGGCTACAGGATCCGCCATCA 933		
4447	TCCGAGCACTTCAGTGGGAGACCTCGAGAAGATCGGTGCCCACTCTCGGAATTC 4503		
934	TCCGAGCACTTCAGTGGGAGACCTCGAGAAGATCGGTGCCCACTCTCGGAATTC 988		

RESULT 11	958 bp	linear	EST 13-MAY-2003
BX416961		mRNA	
LOCUS			
DEFINITION			
	BX416961	Homo sapiens	PLACENTA
	BX416961	Homo sapiens	clone CS0DE002YD03
	5-PRIME	mRNA	sequence.

BX415961
 GI:30654326
 EST.
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (Basse 1 to 958)
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 Contact: Genoscope
 Genoscope - Centre National de Sequençage
 BP 131 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 1413.r For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DE002CE02QP1&cluster=1413.r. Contact :
 Feng Liang Email : fliang@lifetech.com URL : Corporation 1500
 http://fulllength.invitrogen.com/ Invitrogen
 Parady Avenue Genoscope sequence ID : CS0DE002CE02QP1.
 Location/Qualifiers
 1..958
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DE002YD03"
 /issue_type="PLACENTA"
 /clone_lib="Homo sapiens PLACENTA"
 /note="Vector: pCMVSPORT_6; 1st strand cDNA was primed
 with a NotI-oligo(dT) primer. Five prime end enriched,
 double-strand cDNA was digested with Not I and cloned into
 the Not I and EcoRV sites of the pCMVSPORT 6 vector.
 Library was not normalized."

Query Match	11.5%;	Score	895.6;	DB	13;	Length	958;
Best Local Similarity	99.0%;	Pred. No.	2.6e-246;				
Matches	901;	Conservative	0;	Mismatches	-9;	Indels	0;
Gaps	0;						
3894	CCACAACCCCTACAAACGGCCGACAGCAGGAAATCTCTTTGGAAAGAGTGTCCATCTCATC	3953					
49	CCGGNATTCGCGGATCGCGCAGCAGGGAAATCTCTTTGGAAGAGTGTCCATCTCATC	108					
3954	AGAGCTCTCGCATCTTTTGATTAACCTGAGTCCCGGCTCGAGTACAATGTCAGTGTTTACA	4013					
109	AGAGCTCTCGCATCTTTTGATTAACCTGAGTCCCGGCTCGAGTACAATGTCAGTGTTTACA	168					
4014	CTGTCAAGGATGACAAGGAAAGTGTCCTATCTCTGTATACCATCATCCACGCTGTTCCTC	4073					
169	CTGTCAAGGATGACAAGGAAAGTGTCCTATCTCTGTATACCATCATCCACGCTGTTCCTC	228					
4074	CTCCCATGACCTGCGATTCACAACATTTGTCAGACACACATGGTGTTACCTGGGGCTC	4133					
229	CTCCCATGACCTGCGATTCACAACATTTGTCAGACACACATGGTGTTACCTGGGGCTC	288					
4134	CACCCCATCCATTGATTTTAAACCACTTCCTGTCGCTTACTCACCTGTGAAAAATGAGG	4193					
289	CACCCCATCCATTGATTTTAAACCACTTCCTGTCGCTTACTCACCTGTGAAAAATGAGG	348					
4194	AAGATGTTGCAGAGTTGTCAATTTCTCCTTCAGACAATGCGAGTGGTCTTAAACAATCTCC	4253					
349	AAGATGTTGCAGAGTTGTCAATTTCTCCTTCAGACAATGCGAGTGGTCTTAAACAATCTCC	408					
4254	TGCTCTGTACAGAAATGATGATGAGTGTCTCCAGTGTCTACGAAACAACATGAGAGCACAC	4313					
409	TGCTCTGTACAGAAATGATGATGAGTGTCTCCAGTGTCTACGAAACAACATGAGAGCACAC	468					
4314	CTCTTAGAGGAAGACAGAAACAGGCTTTTGATTTCCCACTGGCAATTCGATTTCTGATA	4373					
469	CTCTTAGAGGAAGACAGAAACAGGCTTTTGATTTCCCACTGGCAATTCGATTTCTGATA	528					

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 943)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 1413.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0BA10252G09_CS02375_1&cluster=1413.r.
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0BA10252G09_CS02375_1.
Location/Qualifiers
1. .943
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1065Y122"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

FEATURES
source
11.6%; Score 901.2; DB 13; Length 943;
Best Local Similarity 98.6%; Pred. No. 6.1e-248;
Matches 928; Conservative 0; Mismatches 11; Indels 2; Gaps 2;

ORIGIN
Query Match 11.5%; Score 900.2; DB 9; Length 1201;
Best Local Similarity 98.6%; Pred. No. 6.1e-248;
Matches 928; Conservative 0; Mismatches 11; Indels 2; Gaps 2;

400 ACCTGTGTACAGACAGTGGTGGTCTACTCTGTGGGATGAGTGGCTGAAGACACAG 341
1248 GAAATAAGCAAAATGCTTTGACGTGCTGGGCAACGAGATCAGTCCCAAGAGACAGCTG 1307
340 GAAATAAGCAAAATGCTTTGACGTGCTGGGCAACGAGATCAGTCCCAAGAGACAGCTG 281
1308 TAACCCAGACTTACGGTGGCACTCAATGGAGAGCCATGTCTTACCATTCACCTACA 1367
280 TAACCCAGACTTACGGTGGCACTCAATGGAGAGCCATGTCTTACCATTCACCTACA 221
1368 ATGGCAGGAGCTTCTACTCTCTGCACACAGAGAGGGCGAGACGACGACATCTTTGGTGA 1427
220 ATGGCAGGAGCTTCTACTCTCTGCACACAGAGAGGGCGAGACGACGACATCTTTGGTGA 161
1428 GCACAACTTCGAATATGACGAGGACCAAGAAATACCTCTTCTGCAGACACACACTGTTT 1487
160 GCACAACTTCGAATATGACGAGGACCAAGAAATACCTCTTCTGCAGACACACACTGTTT 101
1488 TGGTTCAGACTCGAGGAGGAAATTCCAATGGTGCCTTGTGCCACTTCCCTTCTCTATACA 1547
100 TGGTTCAGACTCGAGGAGGAAATTCCAATGGTGCCTTGTGCCACTTCCCTTCTCTATACA 41
1548 ACAACCAATTCACACTGATGTCACCTCTGAGGCGCAGAGA 1588
40 ACAACCAATTCACACTGATGTCACCTCTGAGGCGCAGAGA 1

RESULT 10
AL541277
LOCUS
DEFINITION
AL541277 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE005YN01
S-PRIME, mRNA sequence.
ACCESSION
AL541277
VERSION
AL541277.2 GI:30545299
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 15, 2001 this sequence version replaced gi:12872191.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 1413.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DE005CG01QPI&cluster=1413.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DE005CG01QPI.
Location/Qualifiers
1. .1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D8005YN01"
/tissue_type="PLACENTA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo (dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and EcoRV sites of the
pCMVSPORT 6 vector. Library was not normalized."

FEATURES
source
11.5%; Score 900.2; DB 9; Length 1201;
Best Local Similarity 98.6%; Pred. No. 6.1e-248;
Matches 928; Conservative 0; Mismatches 11; Indels 2; Gaps 2;

ORIGIN
Query Match 11.5%; Score 900.2; DB 9; Length 1201;
Best Local Similarity 98.6%; Pred. No. 6.1e-248;
Matches 928; Conservative 0; Mismatches 11; Indels 2; Gaps 2;

400 ACCTGTGTACAGACAGTGGTGGTCTACTCTGTGGGATGAGTGGCTGAAGACACAG 341
1248 GAAATAAGCAAAATGCTTTGACGTGCTGGGCAACGAGATCAGTCCCAAGAGACAGCTG 1307
340 GAAATAAGCAAAATGCTTTGACGTGCTGGGCAACGAGATCAGTCCCAAGAGACAGCTG 281
1308 TAACCCAGACTTACGGTGGCACTCAATGGAGAGCCATGTCTTACCATTCACCTACA 1367
280 TAACCCAGACTTACGGTGGCACTCAATGGAGAGCCATGTCTTACCATTCACCTACA 221
1368 ATGGCAGGAGCTTCTACTCTCTGCACACAGAGAGGGCGAGACGACGACATCTTTGGTGA 1427
220 ATGGCAGGAGCTTCTACTCTCTGCACACAGAGAGGGCGAGACGACGACATCTTTGGTGA 161
1428 GCACAACTTCGAATATGACGAGGACCAAGAAATACCTCTTCTGCAGACACACACTGTTT 1487
160 GCACAACTTCGAATATGACGAGGACCAAGAAATACCTCTTCTGCAGACACACACTGTTT 101
1488 TGGTTCAGACTCGAGGAGGAAATTCCAATGGTGCCTTGTGCCACTTCCCTTCTCTATACA 1547
100 TGGTTCAGACTCGAGGAGGAAATTCCAATGGTGCCTTGTGCCACTTCCCTTCTCTATACA 41
1548 ACAACCAATTCACACTGATGTCACCTCTGAGGCGCAGAGA 1588
40 ACAACCAATTCACACTGATGTCACCTCTGAGGCGCAGAGA 1

```

QY 4775 GGTACATCAGCGGCTTAAACCTGGAGTTGATTATATACCATCACTGTGTATGCTGCTCACT 4834
Db |||||
QY 808 GGTACATCAGCGGCTTAAACCTGGAGTTGATTATATACCATCACTGTGTATGCTGCTCACT 867
Db |||||
QY 4835 GGCGGTGGAGACAGCGCCGCGAGAGCGAAGCAATTTCCATTAAATACCGAACAGAAATT 4894
Db |||||
QY 868 GGCGGTGGAGACAGCGCCGCGAGAGCGAAGCAATTTCCATTAAATACCGAACAGAAATT 927
Db |||||
QY 4895 GACAAACCATCCACATGCAAGTGAAGCCAGTGTTCAGGACCAACAGCATTAGTGTCAAGTGG 4954
Db |||||
QY 928 GACAAACCATCCACATGCAAGTGAAGCCAGTGTTCAGGAC-ACAGCATTAAGTGTCAAGTGG 986
Db |||||
QY 4955 CTGCTTCAAGTTCCTCTTACTGTTTACAGGTAACCAACCACTCCCAAAATAGGACCA 5014
Db |||||
QY 987 CTGCTTCAAGTTCCTCTTACTGTTTACAGGTAACCAACCACTCCCAAAATAGGACCA 1043
Db |||||
QY 5015 GGACCAACCAAACTAAACCTGAGTCCAGTCCAGA 5046
Db |||||
QY 1044 ACMAAAATAAATACTGCGGTGCGATCAACA 1075
Db |||||

RESULT 8
B0688596
LOCUS
DEFINITION B0688596 925 bp mRNA linear EST 15-JUL-2002
AGENCOURT_834:987 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6248842
5', mRNA sequence.
ACCESSION B0688596
VERSION B0688596.1 GI:21813912
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS NIH-MGC http://mgc.nhl.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM2389 row: f column: 11
High quality sequence stop: 659.
Location/Qualifiers
1..925
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6248842"
/tissue_type="ductal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 110"
/note="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN
Query Match 11.6%; Score 905; DB 13; Length 925;
Best Local Similarity 99.4%; Pred. No. 4.8e-249;
Matches 919; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 5160 ACATTGATCGCCCTAAAGGACTGGGATTCACATGATGTGATCGATTCGATCCATCAAAATTG 5219
Db |||||

```

```

1 ACATTGATCGCCCTAAAGGACTGGCATTCACTGATGTGATGTCGATTCCATCAAAATTG 60
5220 CTTGGGAAGCCACACAGGGGCAAGTTTCCAGGTACAGGTGACCTACTCGAGCCCTGAGG 5279
61 CTTGGGAAGCCACACAGGGGCAAGTTTCCAGGTACAGGTGACCTACTCGAGCCCTGAGG 120
5280 ATGGAATCCATGAGCTATTCCCTGCACTGTATGTTGTAAGAACACACTGCAGAGCTGCAAG 5339
121 ATGGAATCCATGAGCTATTCCCTGCACTGTATGTTGTAAGAACACACTGCAGAGCTGCAAG 180
5340 GCCTCAGACGGGTTCTGAGTACACAGTCAAGTGTGTTGCTTGCCTTGACAGTATATGAGA 5399
181 GCCTCAGACGGGTTCTGAGTACACAGTCAAGTGTGTTGCTTGCCTTGACAGTATATGAGA 240
5400 GCCAGGCCCTGTGATTGGAACCCAGTCCACAGCTATTCTCTGCACCACTGACCTGAAGTTCA 5459
241 GCCAGGCCCTGTGATTGGAACCCAGTCCACAGCTATTCTCTGCACCACTGACCTGAAGTTCA 300
5460 CTCAGTCAACCCACAGCCTGAGGCCAGTGGACACCAACCCATGTTCAAGTCACTG 5519
301 CTCAGTCAACCCACAGCCTGAGGCCAGTGGACACCAACCCATGTTCAAGTCACTG 360
5520 GATATCGAGTGGGGTGTGACCCCAAGGAGAACCGGACCAATGAAAGAAATCAACCTTG 5579
361 GATATCGAGTGGGGTGTGACCCCAAGGAGAACCGGACCAATGAAAGAAATCAACCTTG 420
5580 CTCCTGACAGCTCATCCGTGTTGTATCAGGACTTATGTGTGGCCACCAATATGAAGTCA 5639
421 CTCCTGACAGCTCATCCGTGTTGTATCAGGACTTATGTGTGGCCACCAATATGAAGTCA 480
5640 GTGCTGTATGCTCTTAAGGACACTTTGACAAAGACAGCAGCTCAGGGAGTTGTCAACACTC 5699
481 GTGCTGTATGCTCTTAAGGACACTTTGACAAAGACAGCAGCTCAGGGAGTTGTCAACACTC 540
5700 TGGAGAAATGTACGCCCAACAAAGAGGGCTCGTGTGACAGATGCTACTGAGACCAATCA 5759
541 TGGAGAAATGTACGCCCAACAAAGAGGGCTCGTGTGACAGATGCTACTGAGACCAATCA 600
5760 CCATTAGCTGGAGAACCAAGACTGAGACGATCACTGGCTTCCAAAGTTGATGCGTTCCAG 5819
601 CCATTAGCTGGAGAACCAAGACTGAGACGATCACTGGCTTCCAAAGTTGATGCGTTCCAG 660
5820 CCAATGCCAGACTCCAAATCCAGAGAACCAATCAAGCCAGATGTGAGAGCTACCAATCA 5879
661 CCAATGCCAGACTCCAAATCCAGAGAACCAATCAAGCCAGATGTGAGAGCTACCAATCA 720
5880 CTCGCTTACAAACAGGAGCTGACTACAGATCTACTGCTACACCTGATGATGATCAATGCTC 5939
721 CAGGTTTACAAACAGGAGCTGACTACAGATCTACTGCTACACCTGATGATGATGATGATGCTC 780
5940 GGAGCTCCCCCTGTGTGTCATCGAGCGCTCCCACTGCCATTGATGACCAATCCAACTCGTT 5999
781 GGAGCTCCCCCTGTGTGTCATCGAGCGCTCCCACTGCCATTGATGACCAATCCAACTCGTT 840
6000 TCTGTGGCCACCAACCAATTCCTGCTGTATCATGGGCGGCGGCGGCGGCGGCGGCGGCGG 6058
841 TCTGTGGCCACCAACCAATTCCTGCTGTATCATGGGCGGCGGCGGCGGCGGCGGCGGCGG 900
6059 ACCGGCTACATCATCAAGTATGAGA 6083
901 ACCGGCTACATCATCAAGTATGAGA 925

RESULT 9
BX391752/c
LOCUS
DEFINITION BX391752 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1065Y122 3-PRIME, mRNA sequence.
ACCESSION BX391752
VERSION BX391752.1 GI:30607701
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

```

488 GGAAACACATCATCAGATATAATCAACAGTGGGAGCGGACCTACCTAGGCAATGCGTTGTT 547
 494 TGTACTTGTATGGAGAAAGCCGAGGTTTAACTCGAAAGTAAACCTGAAGCTGAAGAG 553
 548 TGTACTTGTATGGAGAAAGCCGAGGTTTAACTCGAGAGTAAACCTGAAGCTGAAGAG 607
 554 ACTTGTCTTTGACAGTACACTGGGACACTTACCGAGTGGGTGACACTTATGAGGCTCT 613
 608 ACTTGTCTTTGACAGTACACTGGGACACTTACCGAGTGGGTGACACTTATGAGGCTCT 667
 614 AAAGACTCCATGATCTGGGACTGTACTCTGCATCGGGGCTGGCGGAGGAGATAAGCTGT 673
 668 AAAGACTCCATGATCTGGGACTGTACTCTGCATCGGGGCTGGCGGAGGAGATAAGCTGT 727
 674 ACCATCGAAACCGCTGCCATGAGAGGGGTGAGTCTCAAGATTGGTGCACACTGGAGG 733
 728 ACCATCGAAACCGCTGCCATGAGAGGGGTGAGTCTCAAGATTGGTGCACACTGGAGG 787
 734 AGACCATGAGACTGGTGTGTACATGTTAGAGTGTGTGTCTTGGTAAATGGAAGGA 793
 788 AGACCATGAGACTGGTGTGTACATGTTAGAGTGTGTGTCTTGGTAAATGGAAGGA 847
 794 GAATGACCTGCAAGCCATAGCTGAGAGTGTGTGTGATCATGCTGCTGGACTTCTTAT 853
 848 GAATGACCTGCAAGCCATAGCTGAGAGTGTGTGTGATCATGCTGCTGGACTTCTTAT 907
 854 GTGGTGGGAAACCGTGGGAGAGCCCTACCAAGGCTGGATGATGATGATGATGATGAT 913
 908 GTGGTGGGAAACCGTGGGAGAGCCCTACCAAGGCTGGATGATGATGATGATGATGAT 967
 914 CTGGGAGAGCGGAGCGCATCTTGCATCTTCTAGAGTGTGTGTGATGATGATGATGATG 973
 968 CTGGGAGAGCGGAGCGCATCTTGCATCTTCTAGAGTGTGTGTGATGATGATGATGATG 1026
 974 ACAAGGACATCTTATAGATTGGAGACACTGGAGCAAGAGGATATCGAGAAACCTG 1033
 1027 ACAGG-----ACATCTATAGATTGGGMACTGGAGCAAGAGATATCGAGAAACCTG 1082
 1034 CTCAGTGCATCTGCACAGCAACCGCCGAGGAGAGTGAAGTGTGAGGAGCAGCACTCT 1093
 1083 GTGCAYVKNAGGACCGCGGGGGTGGAGTGTGRRGMMCTCTKTGACMAYCRGS 1142
 1094 GTCGAGACC 1102
 1143 GRTYTGSCC 1151

RESULT 7
 3X439746
 LOCUS BX439746 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE014YA12
 DEFINITION S-PRIME, mRNA sequence.
 ACCESSION BX439746
 VERSION BX439746.1 GI:30773802
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1. (bases 1 to 1201)
 Li, W.B., Gruber, C., Jesses, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 JOURNAL
 COMMENT
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 1413.r For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DE014BA06QPI&cluster=1413.r. Contact :
 Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Paraday Avenue Genoscope sequence ID : CS0DE014BA06QPI.
 Location/Qualifiers
 1. .1201
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DE014YA12"
 /tissue_type="PLACENTA"
 /clone_lib="Homo sapiens PLACENTA"
 /notes="Vector: pCMVSPORT6; 1st strand cDNA was primed
 with a NotI-cligo(dT) primer. Five prime end enriched,
 double-strand cDNA was digested with Not I and cloned into
 the Not I and EcoRV sites of the pCMVSPORT 6 vector.
 Library was not normalized."

ORIGIN

Query Match 11.9%; Score 924; DB 13; Length 1201;
 Best Local Similarity 94.4%; Pred. No. 1.9e-254;
 Matches 993; Conservative 7; Mismatches 47; Indels 5; Gaps 4;
 3996 ACAATGTCAGTGTTCACATGTCACAGGATGACAAAGAAAGTGTCCCTATCTCTGATACCA 4055
 28 AAAAAACAGCGTGGTACCGTCCGGAATTCGCGGATAGTGTCCCTATCTCTGATACCA 87
 4056 TCATCCAGTGTTCCTCCCTCCCACTGACCTGCGATTCACCAACATTTGTTCCAGACCA 4115
 88 TCATCCAGTGTTCCTCCCTCCCACTGACCTGCGATTCACCAACATTTGTTCCAGACCA 147
 4116 TGCGTGTACCTGGGCTCCACCCCATCCATTCATTAACCAACTTCCTGGTGGTTACT 4175
 148 TGCGTGTACCTGGGCTCCACCCCATCCATTCATTAACCAACTTCCTGGTGGTTACT 207
 4176 CACTGTGAAAAATGAGAAAGATGTTGACAGATGTGCAATTTCTCTTCAGACAATGCGAG 4235
 208 CACTGTGAAAAATGAGAAAGATGTTGACAGATGTGCAATTTCTCTTCAGACAATGCGAG 267
 4236 TGGTCTTAAACAATCTCTCGCTGGTACAGAAATATGATGAGTGTCTCCAGTGTCTACG 4295
 268 TGGTCTTAAACAATCTCTCGCTGGTACAGAAATATGATGAGTGTCTCCAGTGTCTACG 327
 4296 AACAACTGAGAGCACACTCTTTAGAGAAAGACAGAAAAAGTGTCTTGGTATCCCAACTG 4355
 328 AACAACTGAGAGCACACTCTTTAGAGAAAGACAGAAAAAGTGTCTTGGTATCCCAACTG 387
 4356 GCATGTGCTTTCTGATATTACTGCCCACTCTTTTACTGTGCACTGATGCTCTCCGAG 4415
 388 GCATGTGCTTTCTGATATTACTGCCCACTCTTTTACTGTGCACTGATGCTCTCCGAG 447
 4416 CCACCATCACTGGCTACAGGATCCGCCATCATCCGAGCACTTCAGTGGAGACCTCGAG 4475
 448 CCACCATCACTGGCTACAGGATCCGCCATCATCCGAGCACTTCAGTGGAGACCTCGAG 507
 4476 AAGATCGGGTCCCACTCTCGGAATTCCTCATCCCTCACCACTCCTCCTCAGGACAG 4535
 508 AAGATCGGGTCCCACTCTCGGAATTCCTCATCCCTCACCACTCCTCCTCAGGACAG 567
 4536 AGTATGTGCTCAGCATCGTTCCTCTTAATGCGAGAGGAAAGTCCCTTATTTGATGGCC 4595
 568 AGTATGTGCTCAGCATCGTTCCTCTTAATGCGAGAGGAAAGTCCCTTATTTGATGGCC 627
 4596 AACAACTCAAGTGTCTGATGTTCCGAGGAGACCTGGAGTGTGCTGCGACCCCACTCA 4655
 628 AACAACTCAAGTGTCTGATGTTCCGAGGAGACCTGGAGTGTGCTGCGACCCCACTCA 687
 4656 GCCTACTGATCAGCTGGGATGCTCTGCTGTACAGTGAAGTATTTACAGGATCACTTACG 4715
 688 GCCTACTGATCAGCTGGGATGCTCTGCTGTACAGTGAAGTATTTACAGGATCACTTACG 747
 4716 GAGAAACAGGAGGAAATAGCCCTGTCAGGAGTTCAGTGCCTTGGAGCAACTCTACA 4774
 748 GAGAAACAGGAGGAAATAGCCCTGTCAGGAGTTCAGTGCCTTGGAGCAACTCTACA 807

2y	11	CGCCTGTGTGCA	CAGGGGAGAGAGGAA	CCC	CAGGCGCGAGCGGAGAGGGGACCT	70
Db	66	CGCCTGTGTGCA	CAGGGGAGAGAGGAA	CCC	CAGGCGCGAGCGGAGAGGGGACCT	125
Qy	71	GCAGCCACAAC	TTCTCTGTGTCCT	TCGATCCCTT	CTGTCCCTCCACCCGCTCCCCCTTCCCC	130
Db	126	GCAGCCACAAC	TTCTCTGTGTCCT	TCGATCCCTT	CTGTCCCTCCHCCCGTCCCTTTCGCC	185
2y	131	ACCTCTGGCCCC	CACTTCTTGGAGGCGAGCAAC	CCC	CGGAGGCAATTAGAGGGAATTT	190
Db	186	ACCTCTGGCCCC	CACTTCTTGGAGGCGAGCAAC	CCC	CGGAGGCAATTAGAGGGAATTT	245
2y	191	TCCCGCAGGTTG	CGAAGGGAAGCAAA	CTTGTCGCAACTT	GCCTCCCGGTGCGGGCGTCT	250
Db	246	TCCCGCA - GTT	CGAAGGGAAGCAAA	CTTGTCGCAACTT	GCCTCCCGGTGCGGGCGTCT	304
Qy	251	CTCCCCCAGCT	CTCAACATGCTT	AGGGTC	CGGGGCCCGGGCTGCTGCTGTGGCGCTC	310
Db	305	CTCCCCCAGCT	CTCAACATGCTT	AGGGTC	CGGGGCCCGGGCTGCTGCTGTGGCGCTC	364
Qy	311	CTGTGCTGGGGA	CAGCGGTGCCCT	CCACGGGAGCCT	CGAAGAGCAAGAGCGAGGCTCAG	370
Db	365	CAGTGCTGGGA	CAGCGGTGCCCT	CCACGGGAGCCT	CGAAGAGCAAGAGCGAGGCTCAG	424
Qy	371	CAAAATGGTT	CAGCCCCAGT	CCCCGGTGGCTGT	CAGTCAAAAGCAAGCCCGTGTGTTATGAC	430
Db	425	CAAAATGGTT	CAGCCCCAGT	CCCCGGTGGCTGT	CAGTCAAAAGCAAGCCCGTGTGTTATGAC	484
Qy	431	AATGGAAAACA	CTATCAGATAAAT	CAACAGTGGGAGCGGACCT	ACCTAGGCAATGCGTTG	490
Db	485	AATGGAAAACA	CTATCAGATAAAT	CAACAGTGGGAGCGGACCT	ACCTAGGCAATGCGTTG	544
2y	491	GTATTGACTT	GTATGGAGAA	CGCAGGTTTTAA	CTGCGAAAGTAAACCTGAAGCTGAA	550
Db	545	GTATTGACTT	GTATGGAGAA	CGCAGGTTTTAA	CTGCGAGAGTTAACTCGAAGCTGAA	604
Qy	551	GAGACTTGCTT	GACAAGTAGT	CACATGGGAA	CACATTACCGAGTGGGTGACACTTATGAGCGT	610
Db	605	GAGACTTGCTT	GACAAGTAGT	CACATGGGAA	CACATTACCGAGTGGGTGACACTTATGAGCGT	664
Qy	611	CCTAAAGACT	CTCATGTCT	GGGACTGTA	CTGTCATCGGGCTCGGCGAGGGAATAAGC	670
Db	665	CCTAAAGACT	CTCATGTCT	GGGACTGTA	CTGTCATCGGGCTCGGCGAGGGAATAAGC	724
Qy	671	TGTACCAT	CGCAACCGCT	CGCATG	AGGGGGTCACTTACAAGATTGGTGACACTCGG	730
Db	725	TGTACCAT	CGCAACCGCT	CGCATG	AGGGGGTCACTTACAAGATTGGTGACACTCGG	784
Qy	731	AGGAGACCA	CATGAGACT	TGTTGTTACAT	GTATAGAGTGTGTGTCTTGTTAATGAAAA	790
Db	785	AGGAGACCA	CATGAGACT	TGTTGTTACAT	GTATAGAGTGTGTGTCTTGTTAATGAAAA	844
Qy	791	GGAGATGGAC	CTGCAAGCCAT	AGCTGAGAAGT	TGTTTTCATCATGCTGCTGGGACTTCC	850
Db	845	GGAGATGGAC	CTGCAAGCCAT	AGCTGAGAAGT	TGTTTTCATCATGCTGCTGGGACTTCC	904
Qy	851	TATGTGCT	CGGAGAAAC	CGTGGGAAAGCCCT	TACCAAGGCTGGATGATGGTAGATTGPACT	910
Db	905	TATGTGCT	CGGAGAAAC	CGTGGGAAAGCCCT	TACCAAGGCTGGATGATGGTAGATTGPACT	964
Qy	911	TGCTGGGAG	NAAGCGCGGAG	CGCATC	ACTTGTGCAATCTAGAAATAGATCAACAGATCAG	970
Db	965	TGCTGGGAG	NAAGCGCGGAG	CGCATC	ACTTGTGCAATCTAGAAATAGATCAACAGATCAG	1024
Qy	971	GACACAAGGA	CATCCTATAGAA	TTGGAG		998
Db	1025	GACACAAGGA	CATCCTATAGAA	TTGGAG		1052

RESULT 6	LOCUS	1201 bp	linear	EST 13-MAY-2003
BX417945	BX417945	1201 bp	linear	EST 13-MAY-2003
DEFINITION	BX417945	Homo sapiens	PLACENTA	Homo sapiens cDNA clone CS0DE014Y24

5-PRIME, mRNA sequence.
BX417945.1 GI:30642158
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 "(bases 1 to 1201)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
Bp 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 1413.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DE014BG12QP1&cluster=1413.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0DE014BG12QP1.
Location/Qualifiers
1. 1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE014YM24"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-cligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN

Query Match	12.2%;	Score	947.8;	DB 13;	Length	1201;			
Best Local Similarity	91.7%;	Pred.	No. 2.4e-261;						
Matches	999;	Conservative	21;	Mismatches	64;	Indels	5;	Gaps	2;

QY	14	CTGTGCTGCACAGGGGAGGAGAGGAAACCCACAGGCGCGACGCGGAGAGAGGGGACCTGTGCA	73
Db	68	CTGTGCTGCACAGGGGAGGAGAGGAAACCCACAGGCGCGAGCGGAGAGAGGGACCTGTGCA	127
QY	74	GCACAACTTCTCTGTGCTCTGTGCATCCCTTCTGTCCCTCCACCGCTCCCTTCCCCACCC	133
Db	128	GCACAACTTCTCTGTGCTCTGTGCATCCCTTCTGTCCCTTCTTCCGTTCCTTCCCCACCC	187
QY	134	CTCTGCCCCCACCCTTCTTTGAGGCGACAAACCCCGGAGGCATTAGAGGGATTTTTTCC	193
Db	188	CTCTGCCCCCACCCTTCTTTGAGGCGACAAACCCCGGAGGCATTAGAGGGATTTTTTCC	247
QY	194	CGCAGGTTGCGAGGGAAGCAAACTTGTGTGGCAACTTGCTCCCGGTGCGGGCGTCTCTC	253
Db	248	CGCAGGTTGCGAAGGGAAGCAAACTTGTGTGGCAACTTGCTCCCGGTGCGGGCGTCTCTC	307
QY	254	CCCAACGGTCTCAACATGCTTAAGGGTCCGGGGCCCGGGCTGTGCTGTCTGTGGCCGTCTGT	313
Db	308	CCCAACGGTCTCAACATGCTTAAGGGTCCGGGGCCCGGGCTGTGCTGTCTGTGGCCGTCTGT	367
QY	314	TGCTGTGGGACAGCGGTGCCCTCCACGGGAGCCTCGAAGAGCAAGAGGCGAGGCTCAGCAA	373
Db	368	TGCTGTGGGACAGCGGTGCCCTCCACGGGAGCCTCGAAGAGCAAGAGGCGAGGCTCAGCAA	427
QY	374	ATGGTTTCAGCCCGAGTCCCGGGTGTGCAGTCAAGCAAGCCCGGTGTGTTATGACAAT	433
Db	428	ATGGTTTCAGCCCGAGTCCCGGGTGTGTGCTCAAGCAAGCCCGGTGTGTTATGACAAT	487
QY	434	GGAAAAACATTCAGATPAAATCAACAGTGGGAGCGGACCTTACTTAGGCAATGCCGTTGGTT	493

6508	Y	-----AGGACNAGAAGCT 6520
6760	b	TACCACCTCTATCCTCATGTTTCGGGCGCTCAATCCAAATGCGCTTACAGGACAGAAAGCT 6819
6521	y	CTCTCTCAGACAAACCATCTCATGGGCCCCATTCCAGGACACATCTTCGAGTACATCATTTTCA 6580
6820	b	CTCTCTCAGACAAACCATCTCTTGGACGCGGTTCCAGGAGAGTTCTGAGTACATCATTTTCA 6879
6581	y	TGTCATCTCTGTGGCACTGATGAAGAAACCTTTACAGTTTACGGGTTCTCTGGAACATCTCTACC 6640
6880	b	TGCCAACCAAGTTGGCACCGACGACGAGAGCCCTTACAGTTTCCAAAGTTCTTGGAACTTCTCTACC 6939
6641	y	AGTGCCACTCTCAGACGGCTCACAGAGGTGCCACTCAACATCATAGTGGAGGCACATG 6700
6940	b	AGTGGCATCTCTGACTTGGCCCTTACAGAGGGGTCACTTACAAACATCATAGTGGAGGCACATG 6999
6701	y	AAAGACCAAGCAGAGGACATAAGGTTTCGGGAAGAGTTGTTTACCGTGGGCAACTCTCTCAAC 6760
7000	b	CAGAACCCAGAGGAGGCACAAGGTTTCGGGAAGAGTTGTGACTGTGGGCAACGCTGTACAG 7059
6761	y	GGAAGCTTGAACCAACCTTACGATGACTCGTCTGTTGACCCCTCACACAGTGTCCCATTTAT 6820
7060	b	GAAAGCCTTGAACAGCCTTACAGTGACTCATGCTTTGACCCCTTACCGTTTCCCATTTAC 7119
6821	y	GCGGTTTGGAGATGAGTGGGAACCGAATGTCTGAATCAGGCTTTAAACTGTGTGCCAGTGC 6880
7120	b	GCCATTGGAGAGAGTGGGAGCGGTTGTCTGACGCTGGCTTTAAGCTCACATGCCAGTGC 7179
6881	y	TTAGGCTTTTGGAAAGTGGTCAATTTACAGATGTGATTCATCTAGATGTTGCCATCACAATGGT 6940
7180	b	TTGGGCTTTTGGCAGTGGTCAATTTACAGATCGGATTCATCTAAATGGTGGCCATGAACACGT 7239
6941	y	GTGAACCTACAAGATTTGGACAGAAAGTGGGACCGTTCAGGGAGAAAAATGGCCAGATGATGAGC 7000
7240	b	GTCAACTACAAGATCGGAGAGAGTGGGATCGGCAGGGAGAAAAATGGCCAGCGGATGAGC 7299
7001	y	TGCAATGTTCTTGGGAACGGAAAAAGAGAAATTTCAAGTGTGACCCCTCATGAGGCAAGGTGT 7060
7300	b	TGCACATGCTCGGGAAATGGAAAGGGAGAAATTTCAAGTGTGATCCCATGAAGCAACGTCG 7359
7061	y	TACGATGATGGGAAGACATACCAGCTAGAGAAACAGTGGCGAG--AAGGAATATCTCGGTGC 7119
7360	b	TATGACATGGGAAGACCTTACCATGTATGAGNACAGTGGCAGAAAGAAATATCTCGAGC 7419
7120	y	CA--TTTGCTCCTGCACATGCTTTTGGAGGCAGCGGGCTGGCGCTGTGACAACTGCCGCA 7178
7420	b	CACTTTGTCTCCTGCACGCTTTTCGGAGGCGCAGCGGGCTGGCGCTGTGACAACTGCCGCA 7479
7179	y	GACCTGGG---GGTGAACCCAGTCCCGAAGGCGACTACTGSCCAGTGCCTCAACACAGTATT 7235
7480	b	GACCTGGGGTGCTGTGAACCCAGTCCCGATGGAAACCCCGGCCACACTTACACCACTATA 7539
7236	y	CTCAGAGATACCATCAGAGAACAAACACATAATGTTAATTGGCCCAATGAGTGTCTCATGC 7295
7540	b	CACAGAGATACATCAGAGAAACAAACACTAAACGTAAATGGCCCATTTGAGTGTCTCATGC 7599
7296	y	CTTTAGATGTACAGGCTCAGAGAGAGATTTCCGAGAGTAAATCATCTTTTCCAACTCAGA 7355
7600	b	CGCTAGATGTGCAAGCTGCAGAGACGATTTCTCGAGAGTA---ATCTTTCCAGCCCCAC 7655
7356	y	GGAAACAAGCATGTCTCTCTGCCAAGATCCCATCTAAACTGGAGTGTATGTATAGCAGAC---C 7412
7656	b	CCTTACAG--TGTTCTCTCTACCAAGGTCAATCCACACCCAGTGATGTTAGCAGACCCCTC 7713
7413	y	CAGGTTAGAGTCTCTTTCTTTCTTAAAGCCCTTGTCTCTGGAGGAAGTTCTCTCAGGTTTC 7472
7714	b	CATCTTTGAGTGTCTTTTCCACCTTAAAGCCTTTTGTCTCTGGAGGCCATGTTCTCAGGTTTC 7773
7473	y	AGGTCCTCACTCAGCTTCTCTCAAGCATCACGCTG--GGAGTTTCTCTGAGGGGTTTTTCTCATTA 7531
7774	b	AGCACATTTTACAGCTTCTCCAGCATCGCCCGTGGGATGTTTGTAGACTTCTCTCTCCTC 7833
7532	y	AATGAGGCGTGCACATTTGCCCTGTTCTGCTTGGAAAGTATTCATACCGCTCAGTATTTT--- 7589

Db 7834 AATGGTGACAGTTGGTCAACCCCTGTTCTGCTTCAGGGTTTCAGTACTGCTCAGTGTGTTT 7893
 Qy 7590 ----AAATGAAGTGAATCTTAAGATTGTTGGATCAATAGGAAGCATATGCGACCAA 7645
 Db 7894 AAGAGAATCAAAAGTTCCTTATGGTTGGTCTGGGATCAATAGGGAACAACAGGTAGCCAA 7953
 Qy 7646 CCAAGATGCAAAATGTTTTCGAAATGATATGACCAAAATTTTAAGTAGGAAAGTCACCCAAA 7705
 Db 7954 CTAGGA-GGAATGTAATGTAATGCTAGTACCCAAAGACCTTGAGCAGAAAAGTCACCCAGA 8012
 Qy 7706 CACTTCGTCTTTCACCTTAAGTGTCTGCCCGCAATACTGTAGGAACAAG 7754
 Db 8013 CACCTCGTCTTC-TTTTGCCATCTGACCTGCAGCACTGTCAGGACATG 8060

RESULT 3
 AF130095 2067 bp mRNA linear HTC 08-MAY-2001
 LOCUS Homo sapiens clone FLC0562 PRO2841 mRNA, complete cds.
 DEFINITION AF130095
 ACCESSION AF130095
 VERSION AF130095.1 GI:11493493
 KEYWORDS HTC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 2067)
 AUTHORS Zhang,C., Yu,Y., Zhang,S., Wei,H., Bi,J., Zhou,G., Dong,C., Zai,Y.,
 Xu,W., Gao,F., Liu,M. and He,F.
 TITLE Functional prediction of the coding sequences of 75 new genes
 deduced by analysis of cDNA clones from human fetal liver
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 2067)
 AUTHORS Zhang,C., Yu,Y., Zhang,S., Wei,H., Bi,J., Zhou,G., Dong,C., Zai,Y.,
 Xu,W., Gao,F., Liu,M. and He,F.
 TITLE Direct Submission
 JOURNAL Submitted (23-FEB-1999) Department of Experimental Hematology,
 Institute of Radiation Medicine, Beijing Taiping Road 27, Beijing,
 Beijing 100850, P. R. China
 FEATURES
 source
 1. 2067
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="FLC0562"
 /tissue_type="liver"
 /dev_stage="fetus"
 872..1363
 /note="predicted protein of HQ2841"
 /codon_start=1
 /evidence=not experimental
 /product="PRO2841"
 /protein_id="AAG35520.1"
 /db_xref="GI:11493494"
 /translation="MSESGFKLICQLGFGSGHFRCDSSRWCHDNGVNYLIGEKWDEQ
 RGQMMSCCTCLNGKEFEKDPHEATCYDDGKTYHVGQWKEYLGAICSTCFGQ
 GWRNCDNRCPGGEPSPSEGTGGSYNGYSQRYHQRTNTNVCIECFEPMPLDVQADRED
 SRE"

ORIGIN
 Query Match 17.2%; Score 1342.2; DB 11; Length 2067;
 Best Local Similarity 85.5%; Pred. No. 0;
 Matches 1624; Conservative 0; Mismatches 8; Indels 267; Gaps 1;

Qy 5897 ACTGACTACAAAGATCTACCTGTACACCTTGAATGACAATGCTCGAGCTCCCTGTGTGTC 5956
 Db 191 ATTAATTAATAAATCTACCTGTACACCTTGAATGACAATGCTCGAGCTCCCTGTGTGTC 250
 Qy 5957 ATCGACGGCTCCATGCGCATTTGATGACCACTCAACCTGGCTTCTTGGCGCACACACCC 6016
 Db 251 ATCGACGGCTCCATGCGCATTTGATGACCACTCAACCTGGCTTCTTGGCGCACACACCC 310

Db 4540 AGGTCTCGATTCCCAACTGGTTTGGATTCTTCTGATATCACCGCAACTCAITTCAGTGT 4599
2y 4396 GCATCGATTGCTCTCTCGAGCCACCATCACTGGTACAGATCGGCATCATCCGAGCA 4455
Db 4600 CCAGTGGTGGCTCTCTCGGGCCCCCATCACGGGTACATCATCGGCATCACGCGAGCA 4659
2y 4456 CTTCAGTGGGAGACTCGAGAAGATCGGGTGGCCCACTCTCGGAATPCCATCACCTCAC 4515
Db 4660 TTCTGTGGAGACCCAGGAGGATCGAGTGGCCCTCGGGAAITCCATCACCTCAC 4719
2y 4516 CAACCTCACTCCAGGCACAGAGTATGGGTGAGATCGTTGCTCTTAATGGCAGAGGGA 4575
Db 4720 CAACCTTAATCCGGGCACCGAGTACGTTGTGAGATCAITTCGTTAATGGCAGAGGGA 4779
2y 4576 AAGTCCCTTAATGATTGGCCCAACATCAACAGTTTCTGATGTTCCGAGGACACTGGAAGT 4635
Db 4780 GAGCGGCCACTGATGGCCAGCAGCAAGTTTCTGATATTCGAGAGATCTGGAGGT 4839
Qy 4636 TGTGTGCGACCCCAACGCCCTACTGATCAGCTGGGATGCTCTCTGCTGTCAAGTGA 4695
Db 4840 CATTCGCTTCCACCCCAACGAGCTGCTCATCAGTTGGGAAACCCCTGCGTCTCTGTGG 4899
Qy 4696 ATATTACAGGATCACTTACGAGAAACAGAGGAAATAGCCCTGTCAGAGATTCACTGT 4755
Db 4900 CTATTACAGAAATCACTACGAGAGACAGGAGGAAATAGCCCTGTCAGAGATTCACTGT 4959
Qy 4756 GCCTGGGAGCAAGTCTACAGTCAATCCATCAGCGGCTTAAAT - CTTGGAGTTGATPACCA 4814
Db 4960 GCCCGAAGCAAGTCCACAGCCACCATCAACAACTTAAACCCAGGAGCAGACTACACCA 5019
Qy 4815 TCAGTGTATGCTGTCACTGGCGTGGAGACAGCCCGCAGCAGCAAGCCCAATTTCCA 4874
Db 5020 TCACCTGTATGCTGTCACTGGCGTGGGACAGTCCACGAGCAGCAAGCCAGTTTCCA 5079
Qy 4875 TTAATTACCGAACAAGAAATGACAAACCATCCAGATGCAAGTGAACCGATGTTCCAGACA 4934
Db 5080 TCAATTATAAACAAGAAATGACAAGCGCTCCAGATGCAAGTGACAGATGTCACGAGCA 5139
Qy 4935 ACAGATTAGTGTCAAGTGGCTGCTCAAGTCCCTGTTACTGCTTACAGATPACCA 4994
Db 5140 ACAGCATCAGTGTCAAGTGGTGGCTTCAACTTCTCTGTGACAGGCTACAGATGACCA 5199
Qy 4995 CCACCTCCCAAAATGGACAGCAACCAAAATCAAAATGCAAGTCCAGATCAAAACAG 5054
Db 5200 CCACCTCCCAAAATGGCTAGCAACCATCAAAATCAAAATGCAAGTCCAGATCAAAACAG 5259
Qy 5055 AAATGACTATTGAAGCTTGCAGCCCAAGTGGAGTATGTTAGTGTCTATGCTCAGA 5114
Db 5260 AAATGACCAATTGAAGGTTTGCACCACTGTGGAGTACGTTAGTGTATTATGCTCAGA 5319
Qy 5115 ATCCAGCGGAGAGTCAAGCTCTGTTTCAGCTGCAAGTAAACCAATTCATCGCCCTA 5174
Db 5320 ACCGAAACGGAAGAACAGCCCTGTTTCAAACTGAGTGACCAACATTCATGCGCCTA 5379
Qy 5175 AAGGACTGCAATTCATGATGAGATGTCATTCATCAAAATTCCTTGGGAAAGCCCAAC 5234
Db 5380 AAGGACTGCAATTCATGATGAGATGTCATTCATCAAAATTCCTTGGGAAAGCCCAAC 5439
Qy 5235 AGGGCAAGTTTCCAGGTACAGGTGACCTACTCGAGCCCTGAGATGCAATTCATGAGC 5294
Db 5440 AGGGCAAGTTTCCAGGTACAGGTGACCTACTCGAGCCCTGAGATGCAATTCATGAGC 5499
Qy 5295 TATTCCCTGCACTGATGTTGAAGAGACACTGCGAGCTGCAAGCCCTCAGACCGGTT 5354
Db 5500 TTTTCCCTGCACTGATGTTGAAGAGCACTGCGAGAGCTGAGGCGCTCAGGCGGGT 5559
Qy 5355 CTGAGTACAGTGTGTTGCTTGCAGATGATATGAGAGCCAGCCCTGATTG 5414
Db 5560 CTGAGTACAGTGTGTTGCTTGCAGATGATATGAGAGCCAGCCCTGATTG 5619
Qy 5415 GAACCCAGTCCACAGCTATTCTCTGCAACCACTGACCTGAAGTTTCACTCAGGTCAACCCA 5474
Db 5620 GAATCCAGTCCACAGCCATTCTCTGGCCCCCAACTCTGAAGTTTCACTCAGGTGACACCCA 5679

Qy 5475 CAAGCTCGAGCGCCAGTGGACACACCCCAATGTTTCACTCACTGATATCGAGTGGCGG 5534
Db 5680 CCAGCTTTACTGGCCAGTGGATAGACACCACTGTTTCACTCACTGGCTACCGGTGGCGG 5739
Qy 5535 TGACCCCAAGAGAGACCGGACCAATGAAGAAATCAACTTGGTCTCTGACAGCTCAT 5594
Db 5740 TGAAACCCGAAGAGAGACAGACCAATGAAGAAATCAACTTCTCCAGACAGCTCAT 5799
Qy 5595 CGTGTGTTGATCAGACTTATGTTGGCCACCAATATGAAGTGAAGTGTCTATGCTCTTA 5654
Db 5800 CGGTGATTTGTGAGACTCATGTTGGCCACTAATACGAAGTCAAGTGTCTATGCTCTCA 5859
Qy 5655 AGGACATTTGACAGCAGACAGCTCAGGAGTTGTCAACACTCTGGAGAAATGTCAGCC 5714
Db 5860 AGGACACACTGACAGCAGACAGCCAGCCAGGAGTCACTACTCTGGAGAAATGTTAGCC 5919
Qy 5715 CACCAAGAGGCTCGTGTGACAGATGCTACTGAGACCACTACCACTTACTAGCTGAGAA 5774
Db 5920 CTCCAGAAAGGCCCCGTGTGACGAGCTACAGAGCCCACTCACTATTAGCTGAGNA 5979
Qy 5775 CCAAGACTGAGACGATCACTGGCTTCCAAAGTTGATGCGTTCAGCCAAATGGCCAGCTC 5834
Db 5980 CAAAGACAGAGCAATCACTGGCTTCCAAAGTCAAGTCCATCCAGCCAAATGGCCAGACC 6039
Qy 5835 CAATCCAGAAACCATCAAGCCAGATGTCAGAGCTACAGCATCACTGGCTTACAAACAG 5894
Db 6040 CAGTTTCAGAGGAGCTCAGCCCGGATGTTAGAGCTACACCATTAAGGTTTACAGCCAG 6099
Qy 5895 GCATGACTTACAAAGATCTACCTGTATACACTTGAATGACAATGCTCGAGCTCCCTGTGG 5954
Db 6100 GCATGACTTACAAAGATCCACCTGTACACTTGAACGACAATGCCCGAGCTCGCCGTGA 6159
Qy 5955 TCATCGACCCCTCACTGCTGATGATGACCATTCACCTGCTGCTGCTGCTGCTGCTGCTG 6014
Db 6160 TCATCGATCCCTCCACCGCCATCGACGACCATTCACCTGCTGCTGCTGCTGCTGCTGCTG 6219
Qy 6015 CCAATTCCTTGTGTGTATCATGGCAGCCGCAAGTGGCCAGGATTTACCGGCTACATCATCA 6074
Db 6220 CCAATTCCTTGTGTGTATCATGGCAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6279
Qy 6075 AGTATGAGAGCTGGTGTCTCTCCAGAGAAAGTGGTCCCTGCGCCCGCTGCTGCTGCTGCTG 6134
Db 6280 AGTATGAGAGCTGGATCCCTCCAGAGAAAGTGGTCCCTGCGCCCGCTGCTGCTGCTGCTG 6339
Qy 6135 CAGAGCTACTTACTGCTGCTGGA - ACCGGAAACCGAATATACAAATTTATGTCATTTGCC 6193
Db 6340 CGAGGCTCACTTACTGCTGCTGAGGCCAGGAAACCGAGTACACCATCTACGTCAATTGCC 6399
Qy 6194 CTGAAGAAATATCAGAGAGCGGCTGATGGAAGGAAAGACAGACAGCTTCC 6253
Db 6400 CTGAAGAAACAAATCAGAAAGTGAAGCCCTGATTTGGAGGAAAGACAGATGAGCTTCC 6459
Qy 6254 CAACTGTAACCTTCCACACCCCAATCTTCTGAGCAGAGATCTTGGATGTTCTCTTCC 6313
Db 6460 CAACTGTTACCTTCCACACCCCAATCTTCTGAGCAGAGATCTTGGATGTTCTCTTCC 6519
Qy 6314 ACAGTTCAAAAGACCCCTTCTGTCACCCACCTGAGTATGACACTGGAATGATTTAG 6373
Db 6520 ACAGTTCAAAAGACCCCTTCTGTCACCAACCTGAGTATGACACTGGAATGATTTAG 6579
Qy 6374 GTTCTTGGCACTTCTGCTGAGCAACCCAGTGTGGCAACAAATGATCTTTGAGGAACAT 6433
Db 6580 GTTCTTGGCAACCCAGCAACCCAGTGTGGCAACAAATGATCTTTGAGGAACAT 6639
Qy 6434 GGTTTTAGGCGGACCAACCGCCCAACAGCCCACTAGGCAATAGGCAATAGGCAAGCA 6493
Db 6640 GGTTTTAGGCGGACCAACCGCCCAACAGCCCACTAGGCAATAGGCAATAGGCAAGCA 6699
Qy 6494 TACCCCGGAGTGT----- 6507
Db 6700 TACCTGCCGAATGATAGATGAGGAGTCCAAATCGGTGATGTTCCCGAGGAGACGTGGAC 6759

2455 CGAATACAGCCAGTAGCTTTGGTCTCTCCCTGGTCTCAGCTTCGACACCGTGTCCGG 2514
 2380 TGAATACAGCCAGCAGCTTTGGTCTCTCAATGGTCTCAGCTTCGACACCGTGTCCAGG 2439
 2515 ATTCCGGGTGGAATATGACTCAGTCAGGAGGAGATGAGCACAAGTACCTGGATCTTCC 2574
 2440 CTTCCGGGTGGAATATGACTCAGTCAGGAGGAGATGAAACACAGTACCTTGAICTCC 2499
 2575 AGCAGCAGCCACTTCTGTGAACAATCCCTGACCT- GCTTCTGGCCGAAATACATGTAA 2633
 2500 AAGCAGCCGCACTTCCGTGAACAATTCCTGACCTGGCTCCCGGCAAGAGTACATGTCA 2559
 2634 ATGCTATCAGATATCTGAGGATGGGAGCAGAGTTGATCTCTGCTACTTCACAAACA 2693
 2560 ATGCTATCAGATATCTGAGGAGGAAACAGAGCTTGAICTGCTACTTCACAGACTA 2619
 2694 CAGGCTGATGCCCTCTCTGACCCGACTGTGGAACAAGTTGATGACACCTCAATGTG 2753
 2620 CAGCAGCTGACGCTCTCTCAGACCTTACCCTGAGCAGGTTGATGATCTTCCATGTG 2679
 2754 TTCGCTGGAGCAGACCCAGGCTCCCATCACAGGTACAGAAATAGTCTATTCCGCTCAG 2813
 2680 TTCGCTGGAGTAGACCCAGGCACTTATCACAGGTATAGAAATGTCTATTACCTTCAG 2739
 2814 TAGAAGTAGACAGCAGAACTCAACTCTCTGAATCTGAAACTCGAACTCCGTCACCTCAGT 2873
 2874 ACTTGCACCTGGTGTTCAGTATAACATCACTATCTATCTGTGGAAGAAATCAAGAAA 2933
 2800 ACTTGACCCGGTGTTCAGTACAACATCACTATCTATCTGTGGAAGAGAACAGGAGA 2859
 2934 GTACACCTGTTGTCATTCAACAAGAAACCACTGGCAACCCGCTCAGATACAGTGCCT 2993
 2860 GCACCCGTTTTTCATCCAAAGAGACCACTGGCAACCCGCTCAGATCTGATAACCTCC 2919
 2994 CTCAGGAGCTTCAGTTTGTGGAG- TGACAGCTGAGGTCACTCATGTGGACA 3052
 2920 CTCGAGGACCTACAGTTTGTGGAAGTCACTGATGTGAAAGTCACTCATGTGGACC 2979
 3053 CCGCTCAGAGTGCAGTACCGGCTACCGTGTGGATGTGATCCCGTCAACCTGCTGCG 3112
 2980 CCTCTGATGTGTGTCTGATACCGTGTGGAGTCTGCTGCTCAGCTGCGCCGG 3039
 3113 GAGCAGGAGGAGGCTGCCATCAGCAGAAACCTTTGCAAGTCAACCGGCTGCTCC 3172
 3040 GAACATGGCAGAGGCTGCTGTCAACAGAAATACCTTTGCTGAAATCACTGGGCTGCTCC 3099
 3173 CTTGGGCTCACTTATTACTTCAAGTCTTTGCACTG- AGCCATGGGAGGAGAGCAAGCC 3231
 3100 CTTGGGCTCAGTACTCTTCAAGTCTTTGCTGTGAGCAGCCAGGAGGAAAGCAATCC 3159
 3232 TCTGACTGCTCAACAGAACCAACTGAGTGTCTCCACTAACTTCAGTTTGTCAATGA 3291
 3160 TCTGACGGCAACAGACCAACTGACGCTGCCACTAACTTCAGTTTGTCAATGA 3219
 3292 AACTGATTCTACTGCTGTGTGATGAGTCACTTCGAGCTCGGCGCCAGATACAGGATACCG 3351
 3220 AACTGACAGAACAGTTCTGTGTAACGTGACTCCACTTCGAGCCCGTATAGCAGCTACCG 3279
 3352 ACTGACGCTGGGCTTACCCGAGAGAGAGCAGCCAGCAGTACAATGTGGGTCCTCTGT 3411
 3280 ACTGACCGGGGCTGACCCGAGGAGGAGCCAGCCCAAGCAGTACAATGTGGGACCTTGGC 3339
 3412 CTCGAGTACCTGAGGATCTGAGCCTGATCTGAGTACACCTGATCTCCCTCGTGGC 3471
 3340 CTCGAGTATCCCTGAGAAATCTGAGCCTGGTCTGAGTACACCTGACCTTGGTGGC 3399
 3472 CATAAAGGCGCAACAGAGAGCCCAAGCACTGGAGTCTTTACACACTGACGAGCTGG 3531
 3400 TGTGAAGGGAACAGCAGAGTCCCAAGCCACCGAGTCTTTAGTACCTCGCCTGAGCCTCT 3459
 3532 GAGCTCTATTCCACTTACACACCGGAGTCACTGAGACCACTTGTGATCAGATGGAC 3591

3460 GCCTCCATTCCACCTTACACACACCGAGTGCAGAGACCAAAATTTGATCACCCTGGAC 3519
 3592 GCCTGCTCAGAAATTTGTTTAAAGCTGGTGTACACACAGCCAGGAGGAGGAGCACC 3651
 3520 CCGCTCTCAAGGATTTGGCTTCAAGCTGGTGTACACACAGCCAGGAGGAGGAGCACC 3579
 3652 ACAGAAAGTCACTTTCAGACTCAGAAAGCATCGTTGTGCTGGCTTGACTCCAGAGTAGA 3711
 3580 CCGAAGTCACTTTCAGACTCTGGAGCATTTGTTGCTGCTTGACTCCAGCGTGA 3639
 3712 ATAGCTCAACCATCAAGTCTGAGAGATGACAGGAAAGAGATGCCCAATTTGATA 3771
 3640 ATACACTTACACCATCAAGTCTGAGATGCCAGGAGAGATGACCCGATTTGATA 3699
 3772 CAAAGTGGTCAACCAATTTGCTCCACCAAACTTGCATCTGGAGGCAAAACCTGACAC 3831
 3700 CAGTGTAGTCAACCGCTGCTCCACCGACCACTTGCATCTGGAGGCAAAACCTGACAC 3759
 3832 TGGAGTCTCAAGTCTCTGGAGAGGAGCACAACCCAGACATTTACTGGTTATAGAA 3891
 3760 TGGAGTCTTACTGCTCTCTGGAGAGGAGCACTACCCAGATATCAGTGGCTACAGAA 3819
 3892 TACCACAACCCCTACAAACCGCCAGCAGGCAAAATTTCTTGAAGAAAGTGGTCCATGCTGA 3951
 3820 AACTACTACCCCAACGAGCAGGAGGAGGAGCCTCTCTGGAAGAGTGGTCCATGCTGA 3879
 3952 TCAGAGTCTCTGCACTTTTGAATAACCTGAGTCCCGGCTGGAGTACAAATGTCAGTGTTA 4011
 3880 TCAGAGTCTCTGCACTTTTGAAGAACTTCTGGCTGGAGTACAAACGTCAGTGTTA 3939
 4012 CACTGTCAAGGATGACAGGAAAGTGTCCCTATCTCTGTATACCATCATCCCA----- 4063
 3940 CACTGTCAAGATGACAGGAAAGTGTCCCTATCTCTGTATACCGTTGTCCAGAGGTGCC 3999
 4064 ----- 4063
 4000 CCAGTCACTGACCTAAGCTTTGTTGATATAAATGATTCAAGCATCGGCTGAGGTGGAC 4059
 4064 ----- 4063
 4060 CCGCTAACTCTTCCACCATTTATCGGATACCGGATCACAGTAGTTCCGGCAGGAGAGG 4119
 4064 ----- 4063
 4120 GATCCCTATTTTGAAGATTTTGTGGACTCCTCAGCTAGGATACACACAGTTTACAGGG 4179
 4064 ----- 4063
 4180 CTGGAGCCTGSCATTTGATGATCAGCGTTATCACTCTCATTAATGGCGAGAGAT 4239
 4064 -----GCTGTTCTCTCTCCACTGACCTGCGATTCAAC 4096
 4240 GCCCTACTACACTGACACAGCAAAACGCTGTCTCTCCACGGATCTGCGATTCAAC 4299
 4097 AACATTGTCAGACACCATCGTGTACCTGGGCTCCACC-CCATCCATTGATTAAAC 4155
 4300 AATATCGGTCAGACACGATCGGGTCACTTGGGCGCGGCTCCGCTCATCGAGTGCAC 4359
 4156 CAACTTCTGGTGGTACTCACTGCTGAAAATGAGGAAAGATGTCAGAGTTGTCAAT 4215
 4360 CAACTCTTGGTGGTACTCACTCCCGTGAAGATGAAGAGGAGCTTGCAGAGCTATCCAT 4419
 4216 TTCTCTTTCAGCAATGAGTGGTCTTAACTTCTCTCTGCTGGTACAGAAATGTAGT 4275
 4420 TTCACTTTCAGCAATGCGGTCTTAACTTCTCTCTGCTGGGACAGAAATCTTAGT 4479
 4276 GAGTGTCTCCAGTGTCTACGAAACAACTGAGACACCTCTTAGAGGAGGAGGAGGAGG 4335
 4480 CAGTGTCTCCAGTGTCTACGAAACAACTGAGAGCATCTCTCTCCGGGAGGAGGAGGAGG 4539
 4336 AGGTCTTGTATCCCACTGGCATTTCTTGTGATTAATTAAGTCCCACTCTTTTACTGT 4395

Db 168 GCGACTCTCTCCATCCACTCAAGATGCTCAGGGGTCCGGGACCCGGCGGCTGCTGC 227
 Qy 300 TGCTGGCGCTCTGCTGCTGGGACAGCGGTGCCCTCCACGGAGCCCTCGAAGAGCAAGA 359
 Db 228 TGCTGGGAGTCTGCTGCTGGGACCTCGGTGGCTGCTCACCAGAGCGGGAGAGCAAGA 287
 Qy 360 GCGAGGCTCAGCAATGTTGAGCCCGAGTCCCGGGTGGCTGCTCAGTCAAGAGCAAGCCCG 419
 Db 288 GCGAGGCTCAGCAATGCTGAGGCTCAATCCCGGGTGGCTGCTCAGTCAAGAGCAAGCCTG 347
 Qy 420 GTTGTATGACAAATGGAAAAACATATCAGATAAATCAACAGTGGAGCGGAGCTTACCTAG 479
 Db 348 GCTGTTTTGACAAATGGAGACATATCAGATAAATCAGCACTGGAGCGGACCTACCTAG 407
 Qy 480 GCAATGCGTGTGTTGTTACTTGTATGAGAGAGCCGAGGTTTTAACTGCGAAGATTAAC 539
 Db 408 GCAACGCCCTGG--TGTAACCTGCTATGAGGAGAGCCGGGTTTTAACTGCGAGAGCAAGC 465
 Qy 540 CTGAGAGCTGAGAGACTTCTTTGACAAAGTACACTGGGAACACTTACCGAGTGGGTGACA 599
 Db 466 CTGAGCCCTGAAGAGACTTGTCTTTGACAAATACACTGGGAACACTTACAAAGTGGGTGACA 525
 Qy 600 CTATGAGCGTCTTAAAGACTTCATGATCTGGGACTGTACTGCAATCGGGCTGGGCGAG 659
 Db 526 CTATGAGCGCTTAAAGATTCATGATCTGGGACTGTACTGCAATCGGGCTGGGAGAG 585
 Qy 660 GGAGATTAAGCTGTACCATCGCAACCGCTGCCATGAGGGGGTCACTTACCAAGATTG 719
 Db 586 GCAGGATCAGCTGTACCATTGCAAATCGCTGCCATGAGGGGGTCACTTACCAAGATTG 645
 Qy 720 GTGACACCTGGAAGAGACCACTGAGACTGTGTTGATGTTAGAGTGTGTTGTTG 779
 Db 646 GCGACAGTGGAGAGGGCCACATGAGACTGTGTTGATGTTAGAGTGTGTTGTTG 705
 Qy 780 GTAATGGAAGAGAGATGAGACTGCTGAGAGCCCATAGCTGAGAGTGTGTTGATGAGTGTG 839
 Db 706 GAATGGAAGAGAGAGTGAAGCTGCTGAGAGCCCATAGCTGAGAGTGTGTTGATGAGTGTG 765
 Qy 840 CTGGGACTTCTTATGTGCTGGAGAAACCTGGGAGAACCTTACCAAGCGCTGGATGATGG 899
 Db 766 CTGGGAGTCTTACCTGCTGGGAGAGACCTGGGAAGAGCCCTTACCAAGCGCTGGATGATGG 825
 Qy 900 TAGATTGTTACTTCTGCTGGAGAGGAGCGGACCGATCATCTTGCACTTCTAGAAATAGAT 959
 Db 826 TGAGCTGTACTTCTTAGCGAAGGCAATGGAGCGATCATCTGTACTTCCAGAAACAGAT 885
 Qy 960 GCACAGATCAGGACACCAAGGACATCTATAGAAATGGAGACCTTGGAGCAAGAGGATA 1019
 Db 886 GCACAGATCAGGACACCCGACATCTATAGAAATGGAGACCTTGGAGCAAGAGGATA 945
 Qy 1020 ATCGAGGAAACCTGTCTCAGTGATCTGCACAGGCAACCGCGGAGAGTGGAAAGTGTG 1079
 Db 946 ACCGAGAAACCTGTCTCAGTGATCTGCACAGGCAATGGAGAGGGAGTGGAAAGTGTG 1005
 Qy 1080 AGAGGCACACCTGTGTCAGACACATCTGAGCGGATCTGGCCCTTTCACCGATGTTCTGTG 1139
 Db 1006 AGGACACACCTGTGTCAGAAAGTGTTCAGCGGATCTGGCTCTTCTACTGATGTCGAA 1062
 Qy 1140 CAGCTGTTTACCAACCGCAGCTTCAACCCAGCTCTCTCTTATGGCCACTGTGTCAAG 1199
 Db 1063 CAGCTATTTACCAACCGCAGCTTCAACCCAGCTCTCTCTTATGGCCACTGTGTCAAG 1122
 Qy 1200 ACAGTGTGTGTTCTACTGTGGGATGATGCTGAGTGGTGAAGACACAAAGAAATAG-CAA 1258
 Db 1123 ACAGTGTGTGTTCTACTGTGTGGAAATGCAAGTGGTGAAGTGGGAAACAAAGCCAA 1182
 Qy 1259 ATGCTTTGACAGTGTGCTGGGACAGGAGTCACTGTCGCAAGAGACAGCTGTATACCCAGACT 1318
 Db 1183 ATGCTGTGACAGTGTGCTGGGACATGGCTGAGTGGTGGGAGACAGCTGTGACCCAGACT 1242
 Qy 1319 TACGTTGGCACTCAATGAGAGGCC-ATGTTGTTTACATTCACTCAATGGGAGGAC 1377
 Db 1243 TATGGTGGCAATTCAACCGGGAGCCCTGTGTCTCTCCGTTTCACTCAACCGTAGGAC 1302

Qy 1378 GTTCTACTCTCTGACACACAGAAAGGGGAGCAGGACGACATCTTTGGTGCAGACACTTC 1437
 Db 1303 CTTTATTTCTTCCACCCAGAGAGGGCGGCAAGACGACATCTGTGGTGTAGCACAATTC 1362
 Qy 1438 GAATATAGCAGGACCAAGAAATCTCTTTCTGACAGACACACTGTGTTTGGTTGAGAC 1497
 Db 1363 CAATACGAAACAGACCAAGAGTATCTCTTCTGACAGACCAATGCGGTTTGGTTGAGAC 1422
 Qy 1498 TGAGAGAGAAATTCGAATGTTGCTTGTGCCACTTCCCTTCTTATACAAACACCA 1557
 Db 1423 TCAGGCGGAAATTCGAATGTTGCTTGTGCCACTTCCCTTCTTATACAAACCGGAA 1482
 Qy 1558 TTACACTGATGTCACCTTCTGAGGCGAGAGACAAACATGAAGTGTGTGGGACCA-CAC 1616
 Db 1483 TTACACCGACTGTACTTCTGAGGCTCGAGGACCAACATGAATGTTGGCGCACCC 1542
 Qy 1617 AGAACTATGATCCGACCAAGAGTTTGGTCTTGTGCCCTTGTGCCCTTCTTATACAAACACCA 1676
 Db 1543 AGAACTACGATCCGATCAGAAAGTTTGGATTTGTGCCAAATGTTGCCCTTGTGCCCTTGTGCCCT 1602
 Qy 1677 GCACCAACCAATGAAGGGTCTGTACCGCATTTGGAGATCAGTGGGATAAGCAGCATGACA 1736
 Db 1603 GCACCAACCAATGAAGGGTCTGTATCGCATTTGGGATCAGTGGGATAGCAGCATGACC 1662
 Qy 1737 TGGGTCACTGATGAGTGCACGTTGTTGGGAATGTTGTTGGGAAATGTTGTTGGGAAATGTTGTTGG 1796
 Db 1663 TGGGCCACATGATGAGTGCACGTTGTTGGGAAACGTTGTTGGGAAATGTTGTTGGGCTGTCATCC 1722
 Qy 1797 CCTA-CTCCGCGCTTGCAGATCAGTGCATTTGTTGATGATCATCACTTACAAATGTGAACGAC 1855
 Db 1723 CCTAGTCTCCAGCTCCGAGACCAAGTGCATGTTGATGATCACTTACAAATGTGAACGAC 1782
 Qy 1856 ACATTTCCAAAGCGTCAATGAAGAGGGGCACTGCTGAACTGTACATGCTTGGGCTGAGGGT 1915
 Db 1783 ACCTTTCCAAAGCGTCAAGAGGAGGACATATGCTGAACTGTACCTTGTGTTGGTCAAGGC 1842
 Qy 1916 CCGGCGAGTGGAGTGTGATCCGCTGACCAATGCGGAGTTCAGAGATTCAGAGCTGGGAGT 1975
 Db 1843 CCGGCGAGTGGAGTGTGATCCGCTTGAACCAATGCGGAGTTCAGAGATTCAGAGCTGGGAGT 1902
 Qy 1976 TATCAAAATTCGAGATTCATGGGAGAAATGTCATGTTGTCAGATGTCAGATGTCATGCTACTGC 2035
 Db 1903 TACAGATTTGTTGACTCTCTGGGAGAAATTTGTCATGTTGTCGATACCAAGTGTACTGG 1962
 Qy 2036 TATGGCGCTGG-CATTTGGGAGTGGATTCGCAACCTTTACAGACCTTACCAAGCTCAAG 2094
 Db 1963 TACGGCGTGGCCATCGGGAGTGGCACTGTCAACCTCTGCAAGCTTACCGGCAAC 2022
 Qy 2095 TGGTCTGTGGAAGTATTTATCACTGAGACTCCGAGTCAAGCCAACTCCCAACCCCTATCCA 2154
 Db 2023 TGGACCTGTCCAAGTAAATATCAGGAGACCCCGGAGCCCAATTCGCCACCCCTATCCA 2082
 Qy 2155 GTGGAATGCAACAGCCATCTCACAATTTCCAAAGTACATTTCTCAGATGAGAGCTTAAAC 2214
 Db 2083 GTGGAATGCCCCGAGGCTTCAACATCACAAGTACATTTCTCAGATGAGAGCTTAAAC 2142
 Qy 2215 TTCTGTAGGCGGTGGAAGGAGCTACATACAGGCGCACTTAAACTCTTACCAACCTATCAA 2274
 Db 2143 CTCTACGGGTGCTGGAAGGAGCTACATCTCTGGCCACTTAACTCTTATACCACTAA 2202
 Qy 2275 AGGCTTGAAGCGCTGTTGTTATACAGGCGGAGCTCATCAGCATCCAGCATCGGCTACGGCCA 2334
 Db 2203 AGGCTTGAAGCGCTGTTGTTATGAGGAGAGCTCATCAGCATCCAGCATCGGCTATGGCCA 2262
 Qy 2335 CCAGAGTGAAGTGGCTTGTGCTTCAACCCAGCAGCAGCAGCAGCTGTGACCCAGCAA 2394
 Db 2263 CAGAGAGTGAAGCTTGTGCTT-CAACCCAGGCGCAGCAGCAGCTGTGACCCAGCAA 2319
 Qy 2395 CACCGTGAAGGAGAGAGCTCCCTTTCTCTCTTGTGGCCACTTCTGAAATCTGTGAC 2454
 Db 2320 CACCGTGAAGGAGAGAGCTGGCCCTTCTCTCTTGTGGCCACTTCTGAAATCTGTGAC 2379

b 7601 CTGATAGAGACGATTCTCGAGAGTA----ATCTTTCCAGCCCCCCTTACAG--TGCT 7654
y 7371 CTTCTGCCAGATCCATCTAACTGAGTGATCTTAGCAGAC---CCAGCTTAGAGTCTTT 7427
b 7655 CTTCTACCAAGTCAATCCACCCAGATGATGTAGCAGACCTCCATCTTTAGTGGTC 7714
y 7428 CTTTCTTTTAAAGCCCTTGTCTGGAGGAAGTTCTCCAGCTTCAGTCAACTCAGC 7487
b 7715 CTTTCAACCTTAAAGCTTTTGTCTCGAGCCATGTCTCAGCTTCAGCAGACAAATTACAGC 7774
y 7488 TTCTCCAGCATCACCCCTG-GGAGTTTCTGAGGTTTCTCATAAATGAGCGTGCACA 7546
b 7775 TTCTCTAGCATCGCCCTGGATGTTTGTAGACTTCTCTCCCAATGGTGCAGTTGG 7834
y 7547 TTGCTCTGTTCTCTGGAAGTATCAATACCCCTCAGTATTT-----AAATGAAGTGA 7600
b 7835 TCACCCCTGTTCTGCTTCAGGTTTCTAGTACTCTCAGTGTCTTTTAAAGAGATCAAAAGT 7894
y 7601 TTCTAAGATTGTTGGATCAATAGGAAGCATATGAGCAGCAACCAAGATGCAAAATGT 7660
b 7895 TCTTATGTTTGTCTGGATCAATAGGGAACACAGGTAGCAACTAGGA-GGAATGT 7953
y 7661 TTGGAATGATATGACCAAAATTTTAAAGTAGGAAGTACACCAACACTTCTGCTTTCAC 7720
b 7954 ACTGAATGCTAGTACCAAGACCTTCAGCAGGAAGTACCCAGACACTCTGCTTTC-T 8012
y 7721 TTAAGTGTCTGCCCCCATACTAGTGAACAG 7754
b 8013 TTGTCATCTGACCTCGACGACTGTCCAGGACATG 8046

RESULT 2
AK090130
ACUS
DEFINITION
Mus musculus 5 months female bone marrow stroma cell CRL-2028
SR-4987 cDNA, RIKEN full-length enriched library, clone:G431002M24
product:fibronectin 1, full insert sequence.
AK090130
VERSION AK090130.1 GI:26105744
KEYWORDS
SOURCE HTC; CAP trapper.
ORGANISM
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 Carninci, P. and Hayashizaki, Y.
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
2
3 Shibata, K., Itoh, M., Aizawa, K., Nagao, K., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kiteunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Chazara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Osawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 8329)
ADACHI, J., AIZAWA, K., AKIMURA, T., ARAKAWA, T., BONO, H., CARNINCI, P.,
FUKUDA, S., FURUNO, M., HANAGAKI, T., HARA, A., HASHIZUME, W.,
HAYASHIDA, K., HAYATSU, N., HIRAMOTO, K., HIRAO, T., HIROZANE, T.,
HORI, F., ICHIO, K., ISHII, Y., ITOH, M., KAGAWA, I., KASUKAWA, T.,
KATO, H., KAWAI, J., KOJIMA, Y., KONDO, S., KONNO, H., KODA, M.,
KOYA, S., KURIHARA, C., MATSUYAMA, T., MIYAZAKI, A., MURATA, M.,
NAKANURA, M., NISHI, K., NOMURA, K., NUMATA, R., OHNO, M., OHSATO, N.,
OKAZAKI, Y., SAITO, R., SAITO, H., SAKAI, C., SAKAI, K., SAKAZUME, N.,
SANO, H., SASAKI, D., SHIBATA, K., SHINGAWA, A., SHIRAKI, T.,
SOGABE, Y., TAGAMI, M., TAGAWA, A., TAKAHASHI, F., TAKAKU-AKIHARA, S.,
TAKEDA, Y., TANAKA, T., TOMARU, A., TOYA, T., YASUNISHI, A.,
MURAMATSU, M. and HAYASHIZAKI, Y.
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
Location/Qualifiers
1. 8329
/organism="Mus musculus"
/mol_type="mRNA"
/strain="BDF1"
/db_xref="FANTOM,DB:G431002M24"
/db_xref="MGI:2429036"
/db_xref="taxon:10090"
/clone="G431002M24"
/sex="female"
/cell_line="CRL-2028 SR-4987"
/cell_type="stroma cell"
/tissue_type="bone marrow"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="5 months"
194. 7640
/note="fibronectin 1 (MGI:95566, GB|X93167, evidence:
BLASTN, 99%, match=4434)
putative"
8310. 8315
/note="putative"
8329
/note="putative"
ORIGIN
Query Match 66.5%; Score 5187.4; DB 11; Length 8329;
Best Local Similarity 81.7%; Pred. No. 0;
Matches 6509; Conservative 0; Mismatches 1046; Indels 414; Gaps 26;
QY 184 GGAATTTTCCGCGAGTTGCCAAGGAAGCAAACTGGTGGCACTTGCCTCCCGGTGCG 243
DB 108 GCACTTTTCCGCGAGCTGCGAGGAGGCCACAGTTCGTGGCCACTTGCCTCTCTGGGAG 167
QY 244 GCGCTCTCTCC-CCACACGCTCTCAACATGCTTAGGGGTCCGGGGCCCCG---GGCTGTGTC 299

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
COMMENT
FEATURES
source
misc_feature
polyA_signal
polyA_site
ORIGIN
Query Match
Best Local Similarity
Matches
QY
DB
QY

3222 GAACAGTTCTGGTAAACGTGGATCTCACCTCGAGCCGTTATGACAGGCTACCGACTGACCG 3281
3360 TGGGCTTACCCGAAGAGAGACAGCCAGGAGGATACAAATGTGGTCCCTCTGTCTCCAAAGT 3419
3282 CGGCTCTGACCCGAGGAGGCGAGCCCAAGCAGTACAAATGTGGGACCCCTTGGCTCCAAAGT 3341
3420 ACCACTGAGGATCTGACGCTGCACTGAGTACACCCGTTATCCCTCTGGTGGCCATAAAGG 3479
3342 ATCCCTTGAGAAATCTGACGCTTGGTCTGAGTACACCCGTGACCTTGGTGGCTGTGAAAG 3401
3480 GCAACCAAGAGAGCCCAACCAACCTGGAGTCTTTTACCACTGACGCTGGAGCTCTTA 3539
3402 GGAACAGCAGAGTCCCAAGCCACCGAGTCTTTACTACCTCTGACGCTCTGCGCTCCA 3461
3540 TCCACCTTACACACCGAGGTCAGTACACCACTGATGATACATGACATGACGCTGTCTC 3599
3462 TTCACCTTACACACCGAGGTCAGTACACCACTGATGATACATGACATGACGCTGTCTC 3521
3600 CAAGAAATGGTTTAAAGCTGGGTGACAGCAACAGCCAGGAGGAGGACCAACGAGAAAG 3659
3522 CAAGAAATGGTTTAAAGCTGGGTGACAGCAACAGCCAGGAGGAGGACCAACGAGAAAG 3581
3660 TGACTTCAGACTCAGGAGCATGTTGTGTCGCGCTGACTCCAGGAGTAGAATAGTCT 3719
3582 TGACTTCAGACTCAGGAGCATGTTGTGTCGCGCTGACTCCAGGAGTAGAATAGTCT 3641
3720 ACACATCCAAAGTCTGAGAGATGGAAGCAAGAGATGCGCCAAATGTTAAACAAAGTGG 3779
3642 ACACATCCAAAGTCTGAGAGATGGAAGCAAGAGATGCGCCAAATGTTAAACAAAGTGG 3700
3780 TGACACATGTTCTCCACCAACAACTGATCTGGAGGCAACCTGACACTGGAGTGC 3839
3701 TGACACCGCTGTCTCCACCAACAACTGATCTGGAGGCAACCTGACACTGGAGTGC 3760
3840 TCACAGTCTCTGGAGAGGAGCACCACCCAGACATTAAGTGTATGAGAAATTAACACAA 3899
3761 TTAATGTTCTCTGGAGAGGAGCACCACCCAGACATTAAGTGTATGAGAAATTAACACAA 3820
3900 CCCCTACAAAGGCGAGAGGAAATCTTTGGAGAGTGGTCCATGCTGATCAGAGCT 3959
3821 CCCCAAGAGGCGAGAGGAGGACCTCTCTGGAGAGAGTGGTCCATGCTGATCAGAGT 3880
3960 CCTGACCTTTTGAATACTGAGTCCCGGCTGGAGTACAAATGTCAGTGTATACACTGTCA 4019
3881 CCTGACCTTTTGAATACTGAGTCCCGGCTGGAGTACAAATGTCAGTGTATACACTGTCA 3940
4020 AGGATGACAGGAAAGTGTCCCTATCTGATACCATCATCCCA 4063
3941 AAGATGACAGGAAAGTGTCCCTATCTGATACCGTTGTCCAGAGGTCGCCAGCTCA 4000
4064 ----- 4063
4001 CTGACCTTAAGCTTTGTGATATACTGATTAAGCATCGGCTGAGGTGGAGCCCGCTAA 4060
4064 ----- 4063
4061 ACTCTCCACCATTAATCGGTACCGAATCACAGTAGITGCGGAGAGAGGATCCCTA 4120
4064 ----- 4063
4121 TTTTGAAGATTTTGTGAGCTCCTCAGTAGGATAGTACAGGTACAGGCTGGAGCTG 4180
4064 ----- 4063
4181 GCATTGACTATGACATCAGCGTTATCACTCTCAATTAATGGGAGAGAGTGCCTCTACTA 4240
4064 -----GCTGTTCTCTCTCCCACTGACCTCGGATCCCAACAACTGCTC 4106
4241 CACTGACACAGCAACAGGCTGTCCCTCTCCCAAGATCTCGGATTAACCAATATCGTCT 4300
4107 CAGACACATGCTGCTACCTGGGTCACACCCCATCCATTAATTAACCAACTTCTG 4166
4301 CAGACAGGATGCGGCTCACTTGGGCCCCGCTCTCGTCCATCGAGCTGACCAACCTCTTG 4360

QY 4167 TGGCTTACTCACCTGTGAAATAAGAGAGATGTCAGAGTGTCAATTTCTCTTCTAG 4226
Db 4361 TGGCTTACTCACCTGTGAAATAAGAGAGATGTCAGAGTGTCAATTTCTCTTCTAG 4420
QY 4227 ACAATGAGTGGTCTTAAACAAATCTCTGCTGGTACAGAAATATGATGAGTGTCTCCA 4286
Db 4421 ACAATGCGTGGTCTTAAACAAATCTCTGCTGGGACAGAAATCTTAGTCAAGTGTCTCCA 4480
QY 4287 GTGCTACGAAACACATGAGAGACACCTCTTAGAGGAACACAGAAACAGAGTCTTGATT 4346
Db 4481 GTGCTACGAAACACATGAGAGACATCCCTCTCCGGGAAGGACAGAAACAGAGTCTCGATT 4540
QY 4347 CCCAACTGGCAATGACTTTTCTGATATTAATGCGCAACTCTTTTCTGCTGCACTGGATTG 4406
Db 4541 CCCAACTGGTCTTTGATTTCTGATATCAACCGCAACTCAATCACTGTCCACTGGGTGG 4600
QY 4407 CTCCTCGAGCCACCATCACTGGCTACAGGATCGCCCATCATCCGAGCACTTCAGTGGGA 4466
Db 4601 CTCCTCGGCCCCCATCAACCGCTACATCCGCCATCACGCGAGCACTTCGTGCGAA 4660
QY 4467 GACCTCGAGAAAGATCGGGTGCCTCATCTCGGAATTCATCCATCAACCTCACTACTC 4526
Db 4661 GACCCAGGAGATCGAGTGCCTCTCGGCAATTCATCACTCAACCTCACTACTC 4720
QY 4527 CAGGACAGAGTATGCTGTCAGCATCTGCTCTTAATGGCAGAGAGAAAGTCCCTTAT 4586
Db 4721 CCGGACCGAGTATGCTGTCAGCATCTGCTCTTAATGGCAGAGAGAGGCGCCGAC 4780
QY 4587 TGATTTGGCCAAACAATCAACAGTCTCTGATGTCGAGGACCTGCGAAAGTGTGCTCGGA 4646
Db 4781 TGATTTGGCCAAACAAGCCACAGTCTCTGATTTCCGAGAGATCTGGAGTCAATTCCTCCA 4840
QY 4647 CCCCCACAGCCTACTGATCAGCTGGGATGCTCTGTCACAGTGAATATTAACAGGA 4706
Db 4841 CCCCCACAGCCTCTCATCAGTTGGGAAACCCCTCGGCTCTCTGCTGCTATTAACAGGA 4900
QY 4707 TCACCTTACGAGAAACAGGAGGAAATAGCCCTGTCCAGGAGTTCACGTGCTGGAGCA 4766
Db 4901 TCACCTTACGAGAGACAGGAGGAAATAGCCCTGTCCAGGAGTTCACGTGCTGGAGCA 4960
QY 4767 AGTCTACAGCTACCATCAGCGCCCTTAAACCTGAGTGTATTAACCATCACTCTGTATG 4826
Db 4961 AGTCCACAGCCACCATCAACAACTTAAACAGGAGAGACTACCATCACTCTGTATG 5020
QY 4827 CTGCTACCTGGCGGTGAGACAGCCCGCAAGCAGCAAGCCAGTTCATCAATTAATAA 4886
Db 5021 CTGCTACCTGGCGGTGAGAGCAGTCCAGCAAGCAGCAAGCCAGTTCATCAATTAATAA 5080
QY 4887 CAGAAATGACAAACCATCCAGATGCAAGTGAACGATGTCAGGACCAACAGCATCAGT 4946
Db 5081 CAGAAATGACAAAGCCGTCCAGATGCAAGTGAACGATGTCAGGACCAACAGCATCAGT 5140
QY 4947 TCAAGTGGCTGCTTCAAGTTCCTGTTTACGAGTAAACCACTCCCAAAA 5006
Db 5141 TCAAGTGGCTGCTTCACTTCTCTGTCAGGCTACAGAGTACCACTCCCAAAA 5200
QY 5007 ATGACCCAGGACCAACAAATAAACTGAGGTCAGGTCAGATCAACAGAAATGATTTG 5066
Db 5201 ATGGCTTAGGACCATCAAAATAAACTGAGGTCAGGTCAGATCAACAGAAATGATTTG 5260
QY 5067 AAGGCTTGCAGCCCACTGAGTATGTTGTTAGTGTCTATGCTCAGAAATCCAAAGCGAG 5126
Db 5261 AAGGTTTGCACCCACTGTTGAGTACGTTAGTGTATGTTTATGCTCAGAAACCGGACGAG 5320
QY 5127 AGAGTACGCTCTGTTTACAGTCAAGTAAACCAATGATGCTCCCTTAAGGACTGGCAT 5186
Db 5321 AAAGCCAGCCCTGTTTCAAACTGAGTGAACCAATGATGCTCCCTTAAGGACTGGCAT 5380
QY 5187 TCACTGATGATGCTGATTCATCAAAATTTGTTGGAAAGCCCAAGGCGCAAGTTT 5246
Db 5381 TCACTGATGATGCTGATTCATCAAAATTTGTTGGAAAGCCCAAGGCGCAAGTTT 5440

QY	1140	CAGCTGTTTACCAACCGCAGCCTCACCCCGAGCCTCTCTCTCTGCGCAGCTGTGTACAG	1199	Db	2145	CGGCTCGTGAAGGAAGCTACCATTTCTTCTGCGCAGCTTTAACTCTATACCATCAAGGCC	2204
Db	1165	CAGCTATTTTACCAACCGCAGACTCACCCCGAGCCTCTCTCTGCGCAGCTGTGTACCG	1124	QY	2280	TGAAGCCTGTGTGTATACGAGGCGCAGCTCATCAGCATCAGCAGTACGCGCACCAAG	2339
QY	1200	ACAGTGTGTGTGTACTCTCTGTGGGATGCGATGGCTGAAGACACAAAGGAATAAGCAAA	1259	Db	2205	TGACCCAGGTGTGTATCTATGAGGACAGCTCATCAGCATCAGCAGTATGGCCACAG	2264
Db	1125	ACAGTGTGTGTGTCTACTCTGTGGGAATGAGTGGCTGAAGTGGCAAGGAACAAAGCAAA	1184	QY	2340	AAAGTCACTCGCTTTGACTTACCTACCAACAGCAGCAGCAGCAGCAGTGTGACGAGCAACG	2399
QY	1260	TGCTTTGACAGTGGCTGGCAACCGAGTCAAGTCCCAAGAGACAGCTGTAAACCAAGCTT	1319	Db	2265	AAAGTACACAGCTTCGACTT---CACCACAGCGCAGCAGCAGCAGCAGCAGCAGCAGCAG	2321
Db	1185	TGCTGTGACGTGCTGGCAGTAATGGCGTCAAGTCCAGAGACAGCAGCTGACCCAGCTT	1244	QY	2400	TCACAGGAGAGACGACTCCCTTTTCTCTCTGTCGCGCAGCTTCTGAATCTGTGAACGAAA	2459
QY	1320	ACGCTGGCAATCAAAATGGAGAGCAGTGTCTTTACCACTACCACTAACAATGGCAGGACGT	1379	Db	2322	TGACCGGAGAGACTCGGCCCTTACTCTCTGTGTGGCCACTTCTGAACTGTAACTGAAA	2381
Db	1245	ATGCTGGCAATTCAAAACGGGAGCCCTGTGTCTCTCCGCTTCACTCAACAGGTAGACCT	1304	QY	2460	TCACAGCCAGTAGTCTTGTGGTCTCTCTGCGGTCTCGACACACCGTGTGCGGATTC	2519
QY	1380	TCTACTCTGACACACAGAAAGGGCGACAGGAGGACATCTTTGGTCAGACACACTTCGA	1439	Db	2382	TCACAGCCAGCAGCTTGTGGTCTCATGGTCTCAGCCTCGACACCGGTGTGAGGCTTC	2441
Db	1305	TCTATTCTGACACACAGAAAGGGCGGCAAGACGACATCTGTGTGTAGACAACTTCCA	1364	QY	2520	GGGTGGAATATGAGTGTGAGGAGGAGGAGTACGACAGTACCTGTGATCTTCCAAAGCA	2579
QY	1440	ATTATGACGAGACACAGAAATACTCTTTCTGCACAGACCACTGTTTGTTCAGACTC	1499	Db	2442	GGGTGAGTATGAGCTGACGAGGAGGAGGAGTAAACCAAGTACCTGTATCTCCCAAGCA	2501
Db	1365	ATTACGACAGACAGCAAGTATCTCTCTGCACAGACCATGCGGTTTGTTCGGACTC	1424	QY	2580	CAGCCACTTCTGTGAACATCCCTGACCTGCTTCTGCGCGGAGAAAATACATGTAAATGTCT	2639
QY	1500	GAGGAGGAAATTCMAATGGTGGCTGTGCCACTTCCCTTCTCTATACAAACCAACTT	1559	Db	2502	CGGCCACTTCCGTGAACATTCCTGACCTGCTCCGGGAGAGAAAGTACATTTGTCAATGTCT	2561
Db	1425	GAGGCGGAAATTCMAATGGTGTCTGTGCCACTTCCCTTCTGTACAAACCCGGAAT	1484	QY	2640	ATCAGATATCTGAGGATGGGAGCAGAGTTTCTGATCTGTCTACTTCAACAAACAGCGC	2699
QY	1560	ACACTGATTGCACTTCTGAGGCGAGAGAGACAAATGAAGTGTGTGGGACCAACAGACA	1619	Db	2562	ATCAGATATCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	2621
Db	1485	ACACCGACTGTACTTCTGAGGCTCGCAGGAGCAACATGAATGTGTGGGACACCCAGACA	1544	QY	2700	CTGATCCCTCTCTGACCCGAGCTGTGGACCAAGTGTATGACACCTCAATTTGTTTGGCT	2759
QY	1620	ACTATGATCGCGACAGAAAGTTTGGGTTCTGCCCATGCTGCCACAGGAGAAATCTGCA	1679	Db	2622	CTGACCTCTCTCCAGACCTCTACCGTGGACCAAGTGTATGATACTTCCATTTGTTTGGCT	2681
Db	1545	ACTACGATCGCGATCAGAAAGTTTGGATTTCTGCCAATGGCTGCCACAGAGAGATCTGCA	1604	QY	2760	GGAGCAGACCCAGGCTCCCATCAGAGGTACAGAGTGTCTTATTCGCCATCAGTAGAAG	2819
QY	1680	CAACCAATGAAGGGGTCTATGCCCATTTGAGATCAAGTGGGATGAAGCAGCATGACATGG	1739	Db	2682	GGAGTAGACCCAGGACCTATACAGGAGTATAGAAATTTCTATTCACCTTCAGTAGAAG	2741
Db	1605	CAACCAATGAAGGGGTCTATGCCCATTTGAGATCAAGTGGGATGAAGCAGCATGACATGG	1664	QY	2820	GTAGCAGCAGAACTCAACCTTCTCTGAAACTGCAAACTCCCTCACCTCAGTGACTTC	2879
QY	1740	GTCATGATGAGGTGCAAGTGTGTGGGATGCTGTGGGATGGAATGGAATGCTATGCT	1799	Db	2742	GCAAGTACAGAGCTCAACCTCCCTGAAACGGCAGCTCCCTCACCTCAGGACCTGC	2801
Db	1665	GCCATGATGAGGTGCAAGTGTGTGGGATGCTGTGGGATGGAATGGAATGCTATGCT	1724	QY	2880	AACTGGTGTTCAGTATAACATCATCTATCTGTCTGTGGAAGAAATCAAGAAAGTACAC	2939
QY	1800	ACTCCAGCTTCAGATCAGTGCATTTGATGATGATGATGATGATGATGATGATGATGATG	1859	Db	2802	AGCCCGTGTTCAGTACCAATCATCTATCTGTCTGTGAGGAGAAACAGGAGAGCAGC	2861
Db	1725	ACTCCAGCTTCAGATCAGTGCATTTGATGATGATGATGATGATGATGATGATGATGATG	1784	QY	2940	CTGTGTCTATTCAACAAAGAAACCACTGGCAACCCAGCTCAGATACAGTGCCTCTCCCA	2999
QY	1860	TCCACAAGCGTTCATGAAGAGGGGACATGCTGAACTGTACATGCTTTCGGTCAGGCTGGG	1919	Db	2862	CGGACTTACAGTTTGTGGAACCTGACTGATGTGAAGTCAACCATCATGTGGACCCCTCTG	2921
Db	1785	TCCACAAGCGTTCATGAAGAGGGGACATGCTGAACTGTACATGCTTTCGGTCAGGCTGGG	1844	QY	3000	GGGACCTGCAAGTTTGTGGAAGTACAGAGTGTGAGGTCTACCATGTGAGACACCGCTG	3059
QY	1920	GCAAGTGAAGTGTGATCCCTCGACCAATCCAGATTCAGAGTTCAGAGTGGGAGCTTTATC	1979	Db	2922	CGGACCTTACAGTTTGTGGAACCTGACTGATGTGAAGTCAACCATCATGTGGACCCCTCTG	2981
Db	1845	GCAAGTGAAGTGTGATCCCTCGACCAATCCAGATTCAGAGTTCAGAGTGGGAGCTTTATC	1904	QY	3060	AGAGTGCAGTACCGGCTACCGGTGTGATGATCCCGTCAACCTGCGCTGGCGAGCAGC	3119
QY	1980	AAATTCGAGATTCATGGGAGAAAGTGTGATGATGATGATGATGATGATGATGATGATGATG	2039	Db	2982	ATAGTGTGTGTCTGGAATCCGTGTGAGGTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG	3041
Db	1905	AGATTCGATCTCTGGGAGAGTTTGTGATGATGATGATGATGATGATGATGATGATGATG	1964	QY	3120	GGCAGAGGCTGCCCATCAGCAGGAAACCACTTTCGAGAGTCAACCGGCTGTCCCTGGGG	3179
QY	2040	GCGTGGCATTTGGGAGTGGCATTTGCAACCTTTACAGACCTTACAGACCTTACAGTCAAGTGT	2099	Db	3042	GGCAGAGGCTGCCCTGTCAACAGAAATACCTTTGCTGAAATCACTGGGCTGTCCCTGGGG	3101
Db	1965	GCGTGGCATTTGGGAGTGGCATTTGCAACCTTTACAGACCTTACAGACCTTACAGTCAAGTGT	2024	QY	3180	TCACCTATTACTTCAAGTCTTTGCGAGTGGCATTTGGAGGAGAGCAAGCTCTGTGACTG	3239
QY	2100	CTGTCGAAGTATTATCACTGAGACTCCGAGTCCAGCTCAGCCCACTCCACCCCATCTCAGTGA	2159	Db	3102	TCAGTACCTCTTCAAGTCTTTGCGTGTGACCAAGGAGGAGGAGGAGGAGGAGGAGGAGG	3161
Db	2025	CTGTCGAAGTATTATCACTGAGACTCCGAGTCCAGCTCAGCCCACTCCACCCCATCTCAGTGA	2084	QY	3240	CTCAACAGACAAACCAACTGGATGCTCCCACTAACTTCCAGTTTGTCAATGAACTGATT	3299
QY	2160	ATGCAACAGACCATCTCAATTTCAAGTACATTTCTCAGGTGGAGACCTTAAATTTCTG	2219	Db	3162	CACAACAGACCAACCAACTCGACGCTCCCACTAACTTCCAGTTTGTCAATGAACTGACA	3221
Db	2085	ATGCCCCGGAGCTTCAACATCACCAGTACATTTCTCAGTGGAGACCTTAAACCTCTA	2144	QY	3300	CTACTGTCTGTGAGTGTGACTCCACTCGGCCCAAGATTAACAGGATACCGACTGACCG	3359
QY	2220	TAGGCCGTTTGAAGGAGCTTACATACCGGCCACTTAACTCTTACACCACTCAAGGCC	2279				

REFERENCE
AUTHORS

3 Shibata, K., Itoh, M., Aizawa, K., Nagakawa, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitazumi, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujisaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-Format
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

REFERENCE
AUTHORS

4 The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5

REFERENCE
AUTHORS

The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 583-573 (2002)
6 (bases 1 to 8315)

JOURNAL

REFERENCE

AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, M.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ono, M., Oheato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M., and Hayashizaki, Y.

TITLE
JOURNAL

Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Teurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/

FEATURES

source

1. .8315
/organism="Mus musculus"
/mol_type="rRNA"
/strain="BDF1"
/db_xref="FANTOM_DB:G431004D19"
/db_xref="MGI:2429041"
/db_xref="taxon:10090"
/clone="G431004D19"
/sex="female"
/cell_line="CRL-2028 SR-4987"
/cell_type="stroma cell"
/issue_type="bone marrow"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="5 months"
194. .7626
/note="fibronectin 1 (MGD|MGI:95566, GB|X93167, evidence:
BLASTN, 99%, match=4434)
putative"
8296. .8301

misc_feature

184 GGATTTTCCCGCAGGTTGCGAAGGGAAGCAAACTTGGTGGCAACTTGGCTCCCGGTGGG 243
108 GGACTTTTCCCGCAGGTTGCGAAGGGAAGCAAACTTGGTGGCAACTTGGCTCCCGGTGGG 167

polya_signal

244 GGCCTCTCTCC-CCACCTCTCAAATGCTTAGGGTTCGGGGCCCG--GCCTGCTGC 299
168 GGCACCTCTCTCCATCCACTCAAGATGCTCAGGGGTCCGGACCCCGGGCGCTGCTGC 227
300 TGCCTGGCGCTCTGTGCTCGGACACAGCGGTGCCCTCCACGGAGGCTCGAAGAGCAAGA 359
228 TGCCTGGCAGTCTGTGCTCGGACCTCGGTGGCTGCCACCGAAGCCGGGAAGAGCAAGA 287
360 GGCAGGCTCAGCAAAATGGTTTCAGCCCGGTCAGCCCGGTCGTCTAGTCAAGCAAGCCCG 419
288 GGCAGGCTCAGCAAAATGGTTTCAGCCCGGTCAGCCCGGTCGTCTAGTCAAGCAAGCCCG 347
420 GTTCTTATGACAAATGGAACAACTATCAGATAAATCAACAGTGGGAGCGGACTACCTAG 479
348 GCTGTTTGAACAATGGGAAGCACTATCAGATAAATCAGCAGTGGGAACGAGACTACCTAG 407
480 GCAATGGCTGGTTTGTACCTGTTATGGAGGAAGCGAGGTTTAACTCGAAAGTAAAC 539
408 GCAAGCCCTGGTTTGTACCTGTTATGGAGGAAGCGGAGTTTAACTCGAGAGCAAGC 467
540 CTGAAGCTGAAGAGACTGTTTGAACAAGTACACTGGGAACACTTACCAGTGGGTGACA 599
468 CTGAGCCTGAAGAGACTGTTTGAACAAGTACACTGGGAACACTTACCAGTGGGTGACA 527
600 CTATGAGCGCTCTAAAGACTCCATGATCTGGGACTGTACCTGCTCATCGGGGTGGCGAG 659
528 CTATGAGCGCTCTAAAGACTCCATGATCTGGGACTGTACCTGCTCATCGGGGTGGCGAG 587
660 GGAGAAATAGCTGATCAATCGAAAACGCTGCCATGAAGGGGTGAGTCTTCAAGATTG 719
588 CAGGATCAGCTGTACCAATGCAATGCTGCTGCAATGAGGGGTGAGTCTTCAAGATTG 647
720 GTGACACTGAGGAGACCACTGAGACTGGTGGTGTACATGTTAGAGTGTGTGTCTTG 779
648 CCGAAGTGGAGAGAGCCACATGAGACTGGTGGCTACATGTTAGAGTGTGTGTCTTG 707
780 GTATGGAAGAGAGAAATGGACCTGCAAGCCCATAGCTGAGAAAGTGTGATCATGCTG 839
708 GAAATGGAAAAGGGGAATGGACCTGCAAACTATAGCTGAGAAAGTGTGATCATGCTG 767
840 CTGGGACTTCTCTATGCTGGTGGGAGAAACGTTGGGAGAGCCCTACCAAGGCTGGATGAG 899
768 CTGGGACTTCTCTATGCTGGTGGGAGAGCCCTGGGAAAAGCCCTACCAAGGCTGGATGAG 827
900 TAGATTGCTTCTGCTGGAGAGAGGAGCGGAGCATCTTGCATCTTAGAAATAGAT 959
828 TGGACTGTACTTGTCTAGGCGAAGGCAATGGAGCGCATCCTCTTACCTCAGAAACAGAT 887
960 GCAACGATCAGGACACACCGGACATCTTATAGAAATGGAGACACCTGGAGCAAGAGGATA 1019
888 GCAACGATCAGGACACACCGGACATCTTATAGAAATGGAGACACCTGGAGCAAGAGGATA 947
1020 ATCAGGAAACCTGCTCCAGTGCATCTGCACAGGCAACCGCCGAGGAGAGTGGAGTGTG 1079
948 ACCGAGGAAACCTGCTCCAGTGCATCTGCACAGGCAATGGAGAGGAGTGGAGTGTG 1007
1080 AGAGGACACCTGCTGCACAGCAACATCGAGCGGATCTGCCCGCTTCAACGATGTCG 1139
1008 AGCGACA---TGCTCTACAAAGTGTCTCAGCCGATCTGGCTCTTCACTGATGTCGAA 1064

polyA_site
ORIGIN
Query Match 66.3%; Score 5324.8; DB 11; Length 8315;
Best Local Similarity 81.8%; Pred. No. 0;
Matches 6504; Conservative 0; Mismatches 1052; Indels 398; Gaps 15;

/note="putative"
8315
/note="putative"

SUMMARIES

61 GAGGGGACCTGCAGCCACAACCTTCTCTGGTCCCTCTGCATCCCTTCTGTCCCTCCACCCGT 120

Db 61 GA

121 CCGCTTCCCAACCTCTGGCCGCCACCTTCTTGGAGGACAAACCCCGGAGGCATTAG 180
 121 CCGCTTCCCAACCTCTGGCCGCCACCTTCTTGGAGGACAAACCCCGGAGGCATTAG 180
 181 AAGGATTTTCCCGAGGTTGGAGGAAAGCAACTTGGTGGCAACTTGGCTCCCGGT 240
 181 AAGGATTTTCCCGAGGTTGGAGGAAAGCAACTTGGTGGCAACTTGGCTCCCGGT 240
 241 GGGGCGTCTCTCCCCCACCCTCTCAACATGCTTAAAGGGTCCGGGGCCGGGCTCTGCT 300
 241 GGGGCGTCTCTCCCCCACCCTCTCAACATGCTTAAAGGGTCCGGGGCCGGGCTCTGCT 300
 301 GCTGGCGTCTCTGGCTGGGACAGCGGTGCTTCCACGGAGGCTTGAAGAGCAAGAG 360
 301 GCTGGCGTCTCTGGCTGGGACAGCGGTGCTTCCACGGAGGCTTGAAGAGCAAGAG 360
 361 GAGGCTCAGCAAAATGGTTGAGCCCGAGTCCCGGTGGCTGTCACTCAAGCAAGCCCGG 420
 361 GAGGCTCAGCAAAATGGTTGAGCCCGAGTCCCGGTGGCTGTCACTCAAGCAAGCCCGG 420
 421 TTGTTATGACATGGAAACACATTCAGATAAATCAACAGTGGAGCGGACCTACTAGG 480
 421 TTGTTATGACATGGAAACACATTCAGATAAATCAACAGTGGAGCGGACCTACTAGG 480
 481 CAATGGGTTGGTTGTTACTTGTATGGAGAAAGCGAGGTTTAACTGCGAAAGTAAACC 540
 481 CAATGGGTTGGTTGTTACTTGTATGGAGAAAGCGAGGTTTAACTGCGAAAGTAAACC 540
 541 TGAAGCTGAAGAGACTGCTTTGACAAAGTACACTGGGAACACTTACCGAGTGGGTGAC 600
 541 TGAAGCTGAAGAGACTGCTTTGACAAAGTACACTGGGAACACTTACCGAGTGGGTGAC 600
 601 TTATGAGCGTCTTAAAGACTCATGATCTGGAGTGTACTTCCATCGGGCTGGGCGAGG 660
 601 TTATGAGCGTCTTAAAGACTCATGATCTGGAGTGTACTTCCATCGGGCTGGGCGAGG 660
 661 GAGAAATAGCTGTACATCGCAAAACCGCTGCCATGAGGGGTGAGTCTTCAAGATTGG 720
 661 GAGAAATAGCTGTACATCGCAAAACCGCTGCCATGAGGGGTGAGTCTTCAAGATTGG 720
 721 TGACACTGGAGAGACCATGAGCTGGTGTACATGTTAGAGTGTGTGTGTGTGTGTGT 780
 721 TGACACTGGAGAGACCATGAGCTGGTGTACATGTTAGAGTGTGTGTGTGTGTGTGT 780
 781 TAATGAAAGAGAGATGACCTGCAAGCCCATAGCTGAGAAAGTGTGTGTGTGTGTGTGT 840
 781 TAATGAAAGAGAGATGACCTGCAAGCCCATAGCTGAGAAAGTGTGTGTGTGTGTGTGT 840
 841 TGGGACTTCTATGTGGTGGAGAAACGTTGGAGAGCCCTTACCAAGCTGGATGATGGT 900
 841 TGGGACTTCTATGTGGTGGAGAAACGTTGGAGAGCCCTTACCAAGCTGGATGATGGT 900
 901 AGATTGACTTGGCTGGGAGAGGACGCGAGCATCACTTGCATCTTCTAGAAATAGATG 960
 901 AGATTGACTTGGCTGGGAGAGGACGCGAGCATCACTTGCATCTTCTAGAAATAGATG 960
 961 CAACGATCAGGACACAAAGACATCTTATAGATTTGGAGACACTTGGAGCAAGAGATAA 1020
 961 CAACGATCAGGACACAAAGACATCTTATAGATTTGGAGACACTTGGAGCAAGAGATAA 1020
 1021 TGAGGAAACCTGCTCCAGTGCATCTGCACAGGCAACGCGCGAGGAGTGGAGTGTGA 1080
 1021 TGAGGAAACCTGCTCCAGTGCATCTGCACAGGCAACGCGCGAGGAGTGGAGTGTGA 1080
 1081 GAGGACACCTCTGTGTCAGACACATCGAGGAGTCTGGCCCTTCAACGATGTTCGTGC 1140
 1081 GAGGACACCTCTGTGTCAGACACATCGAGGAGTCTGGCCCTTCAACGATGTTCGTGC 1140
 1141 AGCTGTTTACCAACCGAGCTTCAACCCAGGCTCTCTCCCTATGGCACTGTGTACAGA 1200
 1141 AGCTGTTTACCAACCGAGCTTCAACCCAGGCTCTCTCCCTATGGCACTGTGTACAGA 1200
 1201 CAGTGTGTGGTCTACTCTGTGGGGATGTCAGTGGTGAAGACAAAGGAATAAAGCAAT 1260

1201 CAGTGTGTGGTCTACTCTGTGGGATGAGTGGCTGAAGACACAAAGGAATAAAGCAAT 1260
 1261 GCCTTTCAGCTGCTGGGCAACGAGTTCAGCTGCAAGAGACAGCTGTAAACCCAGACTTA 1320
 1261 GCCTTTCAGCTGCTGGGCAACGAGTTCAGCTGCAAGAGACAGCTGTAAACCCAGACTTA 1320
 1321 CGGTGCAACTCAAAATGGAGAGCCATGTGTCTTACCAATTCACCTAATGGCAGGAGCTT 1380
 1321 CGGTGCAACTCAAAATGGAGAGCCATGTGTCTTACCAATTCACCTAATGGCAGGAGCTT 1380
 1381 CTACTCTGACACACAGAGAGGCGAGGAGCGGACATCTTTGGTGCAGCACAATTCGAA 1440
 1381 CTACTCTGACACACAGAGAGGCGAGGAGCGGACATCTTTGGTGCAGCACAATTCGAA 1440
 1441 TTATGAGCAGGACAGAAATACTCTTTTGCACAGACCACTGTTTGGTTCAGACTCG 1500
 1441 TTATGAGCAGGACAGAAATACTCTTTTGCACAGACCACTGTTTGGTTCAGACTCG 1500
 1501 AGAGGAAATTCCTTGGTCTTGGCCCTTGGCCCTTCCCTTCTATACAAACCAATTA 1560
 1501 AGAGGAAATTCCTTGGTCTTGGCCCTTGGCCCTTCCCTTCTATACAAACCAATTA 1560
 1561 CACTGATTCACCTTCTGAGGCGAGAGACAAACATGAAAGTGGTGGGACCAACAGAA 1620
 1561 CACTGATTCACCTTCTGAGGCGAGAGACAAACATGAAAGTGGTGGGACCAACAGAA 1620
 1621 CTATGATCCGACAGAAATTTGGTCTTGGCCCTTGGCCCTTCCCTTCTATACAAACCAATTA 1680
 1621 CTATGATCCGACAGAAATTTGGTCTTGGCCCTTGGCCCTTCCCTTCTATACAAACCAATTA 1680
 1681 AACCAATGAAAGGCTCATGTACCGCATTCAGATTCAGTGGGATTAAGCAGCATGACATGG 1740
 1681 AACCAATGAAAGGCTCATGTACCGCATTCAGATTCAGTGGGATTAAGCAGCATGACATGG 1740
 1741 TCACATGATGAGTGCACGTGTGTGGGAAATGGTGTGGGAAATGGGAAATGGGAAATGGGAA 1800
 1741 TCACATGATGAGTGCACGTGTGTGGGAAATGGTGTGGGAAATGGGAAATGGGAAATGGGAA 1800
 1801 CTGCGAGCTTCAGAGATTCAGTGCATTCATGATGACATCACTTCAATGTGAACGACACAT 1860
 1801 CTGCGAGCTTCAGAGATTCAGTGCATTCATGATGACATCACTTCAATGTGAACGACACAT 1860
 1861 CCAAGAGCTTCATGAGAGGCGGACATGCTGAACTGATGATGATGATGATGATGATGATGAT 1920
 1861 CCAAGAGCTTCATGAGAGGCGGACATGCTGAACTGATGATGATGATGATGATGATGATGAT 1920
 1921 CAGTGGAGTGTGATCCGTCGACCAATGCGAGGATTCAGAGACTGGGACGTTTATCA 1980
 1921 CAGTGGAGTGTGATCCGTCGACCAATGCGAGGATTCAGAGACTGGGACGTTTATCA 1980
 1981 AATTGGAGATTCATGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2040
 1981 AATTGGAGATTCATGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2040
 2041 CCGTGGATTCGGGAGTGGCATTTGCCAACCCTTTACAGACTTACCAAGCTCAAGTGGTCC 2100
 2041 CCGTGGATTCGGGAGTGGCATTTGCCAACCCTTTACAGACTTACCAAGCTCAAGTGGTCC 2100
 2101 TGTGGAAGTATTTTATCACTGAGACTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGT 2160
 2101 TGTGGAAGTATTTTATCACTGAGACTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGT 2160
 2161 TGACACACAGGATTCACATTTCCAGTACATTCAGTGGAGGACCTTAAATTCCTGT 2220
 2161 TGACACACAGGATTCACATTTCCAGTACATTCAGTGGAGGACCTTAAATTCCTGT 2220
 2221 AGGCGGTTGGAAGAGAGTACCATACAGGCGCACTTAAACTCCCTACACCATCAAGGCGCT 2280
 2221 AGGCGGTTGGAAGAGAGTACCATACAGGCGCACTTAAACTCCCTACACCATCAAGGCGCT 2280
 2281 GAAGCTGTGTGGTATATCGAGGCGAGCTCATCAGATCCAGCAGTACGGCCACCAAGA 2340

Db 2281 GAAGCCTGGTGTGTATACGAGGCGCAGCTCATCAGCATCAGCAGTACGGCCACCAAGA 2340
2y AGTGACTCGCTTTTGACTTACCACCAACAGCAGCAGCAGCAGTGTGACCAACACCGT 2400
Db 2341 AGTGACTCGCTTTTGACTTACCACCAACAGCAGCAGCAGCAGTGTGACCAACACCGT 2400
2y GACAGGAGAGAGACTCCCTTTTCTCTCTGTGTGGCCACTTCTGAATCTGTGACCGAAT 2460
Db 2401 GACAGGAGAGAGACTCCCTTTTCTCTCTGTGTGGCCACTTCTGAATCTGTGACCGAAT 2460
2y CACAGCCAGTAGCTTTGTGTGTCTCTGTGGTCTCAGCTTCCGACACCGTGTGCGGATTCG 2520
Db 2461 CACAGCCAGTAGCTTTGTGTGTCTCTGTGGTCTCAGCTTCCGACACCGTGTGCGGATTCG 2520
2y GGTGGATATAGCTGAGTGAGGAGGAGATGAGCCACAGTACCTGATCTTCCAGCAC 2580
Db 2521 GGTGGATATAGCTGAGTGAGGAGGAGATGAGCCACAGTACCTGATCTTCCAGCAC 2580
2y AGCCACTTCTGTGAACATCCCTGACCTGCTTCTGTGGCCGAAATACATTTGTAATGTCTA 2640
Db 2581 AGCCACTTCTGTGAACATCCCTGACCTGCTTCTGTGGCCGAAATACATTTGTAATGTCTA 2640
2y TCAGATATCTGAGATGGGAGCAGAGTTGATCTGTCTACTTTCACAAACAGCGCC 2700
Db 2641 TCAGATATCTGAGATGGGAGCAGAGTTGATCTGTCTACTTTCACAAACAGCGCC 2700
2y TGATGCCCTCTCTGACCGACTGTGGACCAAGTTGATGACACCTCAATTTGTTTCGCTG 2760
Db 2701 TGATGCCCTCTCTGACCGACTGTGGACCAAGTTGATGACACCTCAATTTGTTTCGCTG 2760
2y GAGCAGACCCGAGCTCCATCAAGGGTACAGATAGTCTATTTGGCCATCAGTAGAAG 2820
Db 2761 GAGCAGACCCGAGCTCCATCAAGGGTACAGATAGTCTATTTGGCCATCAGTAGAAG 2820
2y TAGCAGCAGCAATCAACCTTCTGAACTGCAAACTCCGTGACCTCAGTGACTTGCA 2880
Db 2821 TAGCAGCAGCAATCAACCTTCTGAACTGCAAACTCCGTGACCTCAGTGACTTGCA 2880
2y ACCTGGTTCAGTATACATCACTATCTATCTGTGTGGAAGAAATCAAGAAAGTACAC 2940
Db 2881 ACCTGGTTCAGTATACATCACTATCTATCTGTGTGGAAGAAATCAAGAAAGTACAC 2940
2y TGTGTCAATCAAGAAACCACTGGCAGCCACCGCTCAGATACAGTCCCTCCAG 3000
Db 2941 TGTGTCAATCAAGAAACCACTGGCAGCCACCGCTCAGATACAGTCCCTCCAG 3000
2y GGACCTGCAGTTTGTGGAAGTGACAGAGTGAAGGTCAACATCATGTGGACACCGCCTGA 3060
Db 3001 GGACCTGCAGTTTGTGGAAGTGACAGAGTGAAGGTCAACATCATGTGGACACCGCCTGA 3060
2y GAGTGCAGTACCGCTACCGTGTGGATGTGATCCCGTCAACCTGGCGAGCAGCG 3120
Db 3061 GAGTGCAGTACCGCTACCGTGTGGATGTGATCCCGTCAACCTGGCGAGCAGCG 3120
2y GCAGAGGTCGCCATCAGCAGGAACCTTTGAGAAAGTCAACCGGCTGCCCTGGGGT 3180
Db 3121 GCAGAGGTCGCCATCAGCAGGAACCTTTGAGAAAGTCAACCGGCTGCCCTGGGGT 3180
2y CACCTATTACTTCAAGTCTTTGAGTGGCCATGAGGAGGAGCAGAGCCTCTGACTGC 3240
Db 3181 CACCTATTACTTCAAGTCTTTGAGTGGCCATGAGGAGGAGCAGAGCCTCTGACTGC 3240
2y TCAACAGACAACCAACTGGATGCTCCCACTAACTCCAGTTTGTCAATGAAGTATTC 3300
Db 3241 TCAACAGACAACCAACTGGATGCTCCCACTAACTCCAGTTTGTCAATGAAGTATTC 3300
2y TACTGTCTGTGTAGATGGATCTCACTCTGGGGCCAGATACAGGATACGACTGACCGT 3360
Db 3301 TACTGTCTGTGTAGATGGATCTCACTCTGGGGCCAGATACAGGATACGACTGACCGT 3360
2y GGGCTTACCCGAGAGAGACAGCCAGCAGTACAAATGTGGTCCCTCTGTCTCAAGTA 3420
Db 3361 GGGCTTACCCGAGAGAGACAGCCAGCAGTACAAATGTGGTCCCTCTGTCTCAAGTA 3420

Qy 3421 CCACCTGAGGAATCTGACGCTGTGATCTGAGTACACCGTATCCCTCGTGGCCATAAAGGG 3480
Db 3421 CCACCTGAGGAATCTGACGCTGTGATCTGAGTACACCGTATCCCTCGTGGCCATAAAGGG 3480
Qy 3481 CAACCAAGAGAGCCCAAGCCACTGGAGTCTTTTACACACACTGACGCTGGAGCTCTAT 3540
Db 3481 CAACCAAGAGAGCCCAAGCCACTGGAGTCTTTTACACACACTGACGCTGGAGCTCTAT 3540
Qy 3541 TCACCTTACAAACAACGAGGTGACTGAGACACACATTTGTGATCATGAGCGCTGTCTCC 3600
Db 3541 TCACCTTACAAACAACGAGGTGACTGAGACACACATTTGTGATCATGAGCGCTGTCTCC 3600
Qy 3601 AGAATTTGTTTTTAAGCTGGGTGTACGACCAAGCCAGGAGGAGGACACACGAGAAGT 3660
Db 3601 AGAATTTGTTTTTAAGCTGGGTGTACGACCAAGCCAGGAGGAGGACACACGAGAAGT 3660
Qy 3661 GACTTCAGACTCAGAAAGCATCGTTGTGTGCGGCTTGACTCCAGGAGTAGAATACGTCTA 3720
Db 3661 GACTTCAGACTCAGAAAGCATCGTTGTGTGCGGCTTGACTCCAGGAGTAGAATACGTCTA 3720
Qy 3721 CACCATCCAAGTCTGAGAGATGGACAGAAAGAGATGCCAATTTGTAACAAAGTGGT 3780
Db 3721 CACCATCCAAGTCTGAGAGATGGACAGAAAGAGATGCCAATTTGTAACAAAGTGGT 3780
Qy 3781 GACACCATTTGTCTCCACCAACAACTTGCATCTGGAGGCAAAACCTGACACTGGAGTGT 3840
Db 3781 GACACCATTTGTCTCCACCAACAACTTGCATCTGGAGGCAAAACCTGACACTGGAGTGT 3840
Qy 3841 CACAGTCTCTGGGAGAGGAGCACCACCCAGACATTTACTGGTTATAGAAATACCACAC 3900
Db 3841 CACAGTCTCTGGGAGAGGAGCACCACCCAGACATTTACTGGTTATAGAAATACCACAC 3900
Qy 3901 CCTACAAACGGCCAGCAGGGAATTTCTTTGGAAGAGTGTCCAATGCTGATCAGAGTCT 3960
Db 3901 CCTACAAACGGCCAGCAGGGAATTTCTTTGGAAGAGTGTCCAATGCTGATCAGAGTCT 3960
Qy 3961 CTGCACTTTTGTATACCTGAGTCCCGGCTGGAGTACAAATGTCAGTGTTTACACTGTCAA 4020
Db 3961 CTGCACTTTTGTATACCTGAGTCCCGGCTGGAGTACAAATGTCAGTGTTTACACTGTCAA 4020
Qy 4021 GGATGACAAAGAAAGTGTCCCTATCTCTGATACCATCATCCAGCTGTTCTCTCCAC 4080
Db 4021 GGATGACAAAGAAAGTGTCCCTATCTCTGATACCATCATCCAGCTGTTCTCTCCAC 4080
Qy 4081 TGACCTGCAATTCACCAACATTTGTTCCAGACACCAATGCGTGTCACTGGGTCCACCC 4140
Db 4081 TGACCTGCAATTCACCAACATTTGTTCCAGACACCAATGCGTGTCACTGGGTCCACCC 4140
Qy 4141 ATCCATTGATTTTAAACCACTTCTGCTGCTTACTCAGCTGTAACAAATGAGAAAGTGT 4200
Db 4141 ATCCATTGATTTTAAACCACTTCTGCTGCTTACTCAGCTGTAACAAATGAGAAAGTGT 4200
Qy 4201 TGACAGATTTGTCAAATTTCTCTTTCAGACAAATGCAAGTGTGTCTTAAACAAATCTCT 4260
Db 4201 TGACAGATTTGTCAAATTTCTCTTTCAGACAAATGCAAGTGTGTCTTAAACAAATCTCT 4260
Qy 4261 TACAGATATATAGTGTGAGTGTCTCCAGTGTCTTACGAAACAAATGAGAGCAGACCTCT 4320
Db 4261 TACAGATATATAGTGTGAGTGTCTCCAGTGTCTTACGAAACAAATGAGAGCAGACCTCT 4320
Qy 4321 AGGAAGACAGAAACAGAGTCTTGATTTCCCACTGGCAATGCACTTTCTGATATTACTGC 4380
Db 4321 AGGAAGACAGAAACAGAGTCTTGATTTCCCACTGGCAATGCACTTTCTGATATTACTGC 4380
Qy 4381 CAACTCTTTTACTGTGCACTGGATTTCTCTCGAGCCACCATCACTGGCTACAGATCCG 4440
Db 4381 CAACTCTTTTACTGTGCACTGGATTTCTCTCGAGCCACCATCACTGGCTACAGATCCG 4440
Qy 4441 CCATCATCCCGAGCAGCTTCACTGGGAGACCTCGAGAAAGATCGGCTGCCCTCTCGAA 4500
Db 4441 CCATCATCCCGAGCAGCTTCACTGGGAGACCTCGAGAAAGATCGGCTGCCCTCTCGAA 4500

4501 TTCCATCACCCCTCACCACTCCTCAGGCAAGAGTATGTGGTCAGCATCGTTGCTCT 4560
4501 TTCCATCACCCCTCACCACTCCTCAGGCAAGAGTATGTGGTCAGCATCGTTGCTCT 4560
4561 TAATGGCAGAGAGAAAGTCCCTTATGTATGGCCAAACAAATCAACAGTTTCTGATGTTCC 4620
4561 TAATGGCAGAGAGAAAGTCCCTTATGTATGGCCAAACAAATCAACAGTTTCTGATGTTCC 4620
4621 GAGGACCTGGAAGTTGTTGCTCGACCCGCCACAGCCTACTGATCAGCTGGGATGCTCC 4680
4621 GAGGACCTGGAAGTTGTTGCTCGACCCGCCACAGCCTACTGATCAGCTGGGATGCTCC 4680
4681 TGCTGTCACAGTGGAGATATTAAGGATCACTTACGGAGAAAACAGAGGAAATAGCCCTGT 4740
4681 TGCTGTCACAGTGGAGATATTAAGGATCACTTACGGAGAAAACAGAGGAAATAGCCCTGT 4740
4741 CCAGGAGTTCACTGCTGGAGCAAGTCTCAGCTACCATCAGCGGCTTAAACCTGG 4800
4741 CCAGGAGTTCACTGCTGGAGCAAGTCTCAGCTACCATCAGCGGCTTAAACCTGG 4800
4801 AGTTGATTAACCATCACTGTATGCTGTCATGCGCTGCGTGGAGACAGCCCGCAAGCAG 4860
4801 AGTTGATTAACCATCACTGTATGCTGTCATGCGCTGCGTGGAGACAGCCCGCAAGCAG 4860
4861 CAAGCCAAATTCATTAATTAACCGAACAGAAATGACAAACATCCAGATGCCAGTGAC 4920
4861 CAAGCCAAATTCATTAATTAACCGAACAGAAATGACAAACATCCAGATGCCAGTGAC 4920
4921 CGATGTTCCAGCAACAGCATTAAGTCAAGTGGCTGCTTCAAGTTCCCTGTTACTGG 4980
4921 CGATGTTCCAGCAACAGCATTAAGTCAAGTGGCTGCTTCAAGTTCCCTGTTACTGG 4980
4981 TTACAGAGTAACACCACTCCCAAAATGACAGGACCAACAAACAACTAAACTCCAG 5040
4981 TTACAGAGTAACACCACTCCCAAAATGACAGGACCAACAAACAACTAAACTCCAG 5040
5041 TCCAGATCAACAGAAATGACTATTAAGGCTTTCAGGCCACAGTGGAGTATGTGTTAG 5100
5041 TCCAGATCAACAGAAATGACTATTAAGGCTTTCAGGCCACAGTGGAGTATGTGTTAG 5100
5101 TGCTATGCTCAGATCCAGCGGAGAGTTCAGGTCAGCTGCTGTTTACAGATCCGATACCAA 5160
5101 TGCTATGCTCAGATCCAGCGGAGAGTTCAGGTCAGCTGCTGTTTACAGATCCGATACCAA 5160
5161 CAATTGATCGCCCTAAAGGACTGGCAATCACTGATGTGGATGTGATTCATCAAAATTCG 5220
5161 CAATTGATCGCCCTAAAGGACTGGCAATCACTGATGTGGATGTGATTCATCAAAATTCG 5220
5221 TTGGGAAAGCCCAAGGGGCAAGTTTCCAGGTCAGGTCAGCTGCTGCTGCTGCTGCTGCTG 5280
5221 TTGGGAAAGCCCAAGGGGCAAGTTTCCAGGTCAGGTCAGCTGCTGCTGCTGCTGCTGCTG 5280
5281 TGGAAATCCATGAGCTATTCCTGCACTGATGTGAAGAGACACTGCAAGCTGCAAGG 5340
5281 TGGAAATCCATGAGCTATTCCTGCACTGATGTGAAGAGACACTGCAAGCTGCAAGG 5340
5341 CCTCAGACCGGGTCTGAGTACACAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5400
5341 CCTCAGACCGGGTCTGAGTACACAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5400
5401 CCAGCCCTGATTCGGAACCCAGTCCAGCTATTCCTGCAACCACTGACCTGAGTTTCA 5460
5401 CCAGCCCTGATTCGGAACCCAGTCCAGCTATTCCTGCAACCACTGACCTGAGTTTCA 5460
5461 TCAGGTCAACCCCAAGCTGAGCGCCAGTGAACACCAACCAATTTTTCAGTCTCATCTGG 5520
5461 TCAGGTCAACCCCAAGCTGAGCGCCAGTGAACACCAACCAATTTTTCAGTCTCATCTGG 5520
5521 ATATCGAGTGGGTGACCCCAAGGAGAGACCGGACCAATGAAGAAATCAACCTTGC 5580
5521 ATATCGAGTGGGTGACCCCAAGGAGAGACCGGACCAATGAAGAAATCAACCTTGC 5580
5581 TCCTGACAGCTCATCCGTGGTGTGATCAGGACTTATGGTGGCCCAACCAATATGAAGTGAG 5640

5581 TCCTGACAGCTCATCCGTGGTGTGATCAGGACTATGGTGGCCCAACAAATATGAAGTGAG 5640
5641 TGCTATGCTCTTAAGGACACTTTTGACAGCAGACAGCTCAGGAGTGTGTCAACACTCT 5700
5641 TGCTATGCTCTTAAGGACACTTTTGACAGCAGACAGCTCAGGAGTGTGTCAACACTCT 5700
5701 GGAGAAATGTGAGCCCAACAAAGAGGGCTCGTGTGACAGATGCTACTGAGACCAACATCAC 5760
5701 GGAGAAATGTGAGCCCAACAAAGAGGGCTCGTGTGACAGATGCTACTGAGACCAACATCAC 5760
5761 CATTAGCTGGAGACCAAGACTGAGACGATCACTGGCTCCAAAGTTGATCCGCTTCCAGC 5820
5761 CATTAGCTGGAGACCAAGACTGAGACGATCACTGGCTCCAAAGTTGATCCGCTTCCAGC 5820
5821 CAATGGCCAGACTCCAAATCCAGAGAACCATCAAGCCAGATGTGAGAGCTPACACCATCAC 5880
5821 CAATGGCCAGACTCCAAATCCAGAGAACCATCAAGCCAGATGTGAGAGCTPACACCATCAC 5880
5881 TGGCTTACAAACCGGCACTGACTACAGAGATCACTGTACACCTTGAATGACATGCTCG 5940
5881 TGGCTTACAAACCGGCACTGACTACAGAGATCACTGTACACCTTGAATGACATGCTCG 5940
5941 GAGCTCCCTCTGTGTCATCGACGCTCCACTGCTGCTCCATGATGACCAATCCAACTGCTG 6000
5941 GAGCTCCCTCTGTGTCATCGACGCTCCACTGCTGCTCCATGATGACCAATCCAACTGCTG 6000
6001 CCTGGCCACCAACCCAAATTCCTTGTGCTGATCATGCGACGCGCCACGTCGCCAGTATAC 6060
6001 CCTGGCCACCAACCCAAATTCCTTGTGCTGATCATGCGACGCGCCACGTCGCCAGTATAC 6060
6061 CGGCTACATCATCAAGTATGAGAAAGCTGGTCTCTCTCCAGAGAAAGTGTCTCCCTCGGCC 6120
6061 CGGCTACATCATCAAGTATGAGAAAGCTGGTCTCTCTCCAGAGAAAGTGTCTCCCTCGGCC 6120
6121 CGGCTTGTGTGACAGAGGCTACTATGCTGCTGAGACCCGGAACCGGAATATCAAT 6180
6121 CGGCTTGTGTGACAGAGGCTACTATGCTGCTGAGACCCGGAACCGGAATATCAAT 6180
6181 TTATGTCATTGCCCTGAAGAATAATCAGAGAGCGAGCCCTGATTGGAAGGAAAAAGAC 6240
6181 TTATGTCATTGCCCTGAAGAATAATCAGAGAGCGAGCCCTGATTGGAAGGAAAAAGAC 6240
6241 AGACAGGCTTCCCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6300
6241 AGACAGGCTTCCCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6300
6301 GGAATGCTTCCCAAGTTCAGAAAGACCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6360
6301 GGAATGCTTCCCAAGTTCAGAAAGACCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6360
6361 AAATGGTATTCAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6420
6361 AAATGGTATTCAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6420
6421 CTTTGGAGAAATGTTTGGCGGACACACCCGCAACCGCCACCGCCACCGCCACCGCCACCGCC 6480
6421 CTTTGGAGAAATGTTTGGCGGACACACCCGCAACCGCCACCGCCACCGCCACCGCCACCGCC 6480
6481 TAGGCAAGACCATACCCGCGAATGTAGGACAAAGAGCTCTCTCTCAGACCAACATCTC 6540
6481 TAGGCAAGACCATACCCGCGAATGTAGGACAAAGAGCTCTCTCTCAGACCAACATCTC 6540
6541 ATGGGCCCCATTCAGGACACTTCTGAGTACATATTTCTGCTGCTGCTGCTGCTGCTGCTG 6600
6541 ATGGGCCCCATTCAGGACACTTCTGAGTACATATTTCTGCTGCTGCTGCTGCTGCTGCTG 6600
6601 TGAAGAACCTTACAGTTTCAAGTTTCTGGAATTTCTACAGTCCCACTCTGTGACAGGCT 6660
6601 TGAAGAACCTTACAGTTTCAAGTTTCTGGAATTTCTACAGTCCCACTCTGTGACAGGCT 6660
6661 CACAGAGGTGCGCCACTCAACATCATGATGAGGACTGAAAGACCAAGCAGCAGAGGCAATA 6720

Db	6661	CACGAGGTGCACCTTACAACATCATAGTGGAGGCMCTGAAGA	CCAGCAGAGGCATAA	6722
Qy	6721	GGTTCCGGAGAGAGGTTGTTACCGTGGCAACTCTGTCAACG	AGAGGCTTGAAACCAACTTAC	6780
Db	6721	GGTTCCGGAGAGAGGTTGTTACCGTGGCAACTCTGTCAACG	AGAGGCTTGAAACCAACTTAC	6780
Qy	6781	GGATGACTCGTGTGTTGACCCCTTACACAGTGTCCCATTT	ATGCGCGTTGGAGATCAGTGGGA	6840
Db	6781	GGATGACTCGTGTGTTGACCCCTTACACAGTGTCCCATTT	ATGCGCGTTGGAGATCAGTGGGA	6840
Qy	6841	ACGAATGTCGTAATTCAGGCCTTTTAAACHTGTTGTCAGTG	CTTATAGGCTTTGGAAAGTGGTCA	6900
Db	6841	ACGAATGTCGTAATTCAGGCCTTTTAAACHTGTTGTCAGTG	CTTATAGGCTTTGGAAAGTGGTCA	6900
Qy	6901	TTTTCAGATGTGAATTCATCTAGATGGTGCCATGACAAATG	GTGTGAACTTACAAGATTTGGAGA	6960
Db	6901	TTTTCAGATGTGAATTCATCTAGATGGTGCCATGACAAATG	GTGTGAACTTACAAGATTTGGAGA	6960
Qy	6961	GAAGTGGGACCGTTCAGGGAGAAATGCGCAGATGATGAGCT	GTGCACATGCTCTGGGACCG	7020
Db	6961	GAAGTGGGACCGTTCAGGGAGAAATGCGCAGATGATGAGCT	GTGCACATGCTCTGGGACCG	7020
Qy	7021	AAAAGGAGAAATTCAAAGTGTGACCCCTCATAGAGGCAACG	TGTATCGATATGGGAAGACATA	7080
Db	7021	AAAAGGAGAAATTCAAAGTGTGACCCCTCATAGAGGCAACG	TGTATCGATATGGGAAGACATA	7080
Qy	7081	CCAGTATGGAGAACAGTGGCAGAGGAATATCTCGGTGCCAT	TTTGTCTGCATCATGCTT	7140
Db	7081	CCAGTATGGAGAACAGTGGCAGAGGAATATCTCGGTGCCAT	TTTGTCTGCATCATGCTT	7140
Qy	7141	TGGAGGCCACGGGGCTTGGCGCTGTGACAACTGCCGCAGAC	CTTGGGGGTGAAACCCAGTCC	7200
Db	7141	TGGAGGCCACGGGGCTTGGCGCTGTGACAACTGCCGCAGAC	CTTGGGGGTGAAACCCAGTCC	7200
Qy	7201	CGAAGGCACACTGTCGCGAGTCTTCAACACAGTATTTCTC	AGAGATACCATTACAGAAACAA	7260
Db	7201	CGAAGGCACACTGTCGCGAGTCTTCAACACAGTATTTCTC	AGAGATACCATTACAGAAACAA	7260
Qy	7261	CACATAATGTTAAATTCGCCCAATTTAGTGTGCTTCATGCC	TTTATAGATCAGAGCTGCACAGAGA	7320
Db	7261	CACATAATGTTAAATTTGCCCAATTTAGTGTGCTTCATGCC	TTTATAGATCAGAGCTGCACAGAGA	7320
Qy	7321	AGATTCGGAGAGTAATTCATCTTTCCAAATCCAGAGGAACA	AGAGCATGCTCTCTGCGCAAG	7380
Db	7321	AGATTCGGAGAGTAATTCATCTTTCCAAATCCAGAGGAACA	AGAGCATGCTCTCTGCGCAAG	7380
Qy	7381	ATCCATCTAAACTCGAGTGATGTTAGCAGACCCAGCTTAGA	GTTCCTTTCTTTCTTAA	7440
Db	7381	ATCCATCTAAACTCGAGTGATGTTAGCAGACCCAGCTTAGA	GTTCCTTTCTTTCTTAA	7440
Qy	7441	GCCCTTTGCTCTCGAGGAGTCTCCAGCTCAGCTCAACTCA	CACAGCTTCTCCCAAGATC	7500
Db	7441	GCCCTTTGCTCTCGAGGAGTCTCCAGCTCAGCTCAACTCA	CACAGCTTCTCCCAAGATC	7500
Qy	7501	ACCTGGGAGTTTCCTGAGGGTTTTCTCATAAATGAGGGCT	GCACATGCGCTGTTCTGCT	7560
Db	7501	ACCTGGGAGTTTCCTGAGGGTTTTCTCATAAATGAGGGCT	GCACATGCGCTGTTCTGCT	7560
Qy	7561	TCGAAGTATTCAAATACCGCTCAGTATTTTAAATCAGGTG	ATTTCTTAAAGATTTGGTTGGGA	7620
Db	7561	TCGAAGTATTCAAATACCGCTCAGTATTTTAAATCAGGTG	ATTTCTTAAAGATTTGGTTGGGA	7620
Qy	7621	TCAATAGGAAAGCATATGCGCCCAACCAAGATGCAAAATG	TTTTGAAATGATATGACCCAAA	7680
Db	7621	TCAATAGGAAAGCATATGCGCCCAACCAAGATGCAAAATG	TTTTGAAATGATATGACCCAAA	7680
Qy	7681	ATTTTAAAGTAGGAAAGTCACCCAAACACTTCTGCTTTT	TCACATTAAGTGTCTGGCCGCAAT	7740
Db	7681	ATTTTAAAGTAGGAAAGTCACCCAAACACTTCTGCTTTT	TCACATTAAGTGTCTGGCCGCAAT	7740
Qy	7741	ACTGTAGGAAACAAGCATGTTGTTACTGTGATATTTTAAAT	TATATCCACAGTACT	7795
Db	7741	ACTGTAGGAAACAAGCATGTTGTTACTGTGATATTTTAAAT	TATATCCACAGTACT	7795

```

RESULT 2
US-10-240-965-121
; Sequence 121, Application US/10240965
; Publication No. US20030165924A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: SHIFFMAN, Dov
; APPLICANT: SOMOGYI, Roland
; APPLICANT: LAWN, Richard M.
; APPLICANT: SEITHAMER, Jeffrey J.
; APPLICANT: PORTER, Gordon J.
; APPLICANT: MIKITA, Thomas
; APPLICANT: TAI, Julie
; TITLE OF INVENTION: GENES EXPRESSED IN FOAM CELL DIFFERENTIATION
; FILE REFERENCE: PA-0025 PCT
; CURRENT APPLICATION NUMBER: US/10/240,965
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: 60/195,106
; PRIOR FILING DATE: 2000-04-05
; NUMBER OF SEQ ID NOS: 276
; SOFTWARE: PERL Program
; SEQ ID NO 121
; LENGTH: 8044
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030165924A1 427813.14
US-10-240-965-121

```

Query Match	99.8%; Score 7780.6; DB 14; Length 8044;
Best Local Similarity	99.9%; Pred. No. 0;
Matches 7786; Conservative	0; Mismatches 9; Indels 0; Gaps 0;
QY	1 CGGCTCAGACCGGTGTGCTSCACAGGGGAGGAGAGGAACCCAGCGCGAGCGGAA 60
DB	
QY	1 CGGCTCAGACCGGTGTGCTSCACAGGGGAGGAGAGGAACCCAGCGCGAGCGGAA 60
DB	
QY	61 GAGGGGACCTGCAGCCCAAACTTCTGTGTCCTGTGCATCCCTTCTGTCTCCACCGGT 120
DB	
QY	61 GAGGGGACCTGCAGCCCAAACTTCTGTGTCCTGTGCATCCCTTCTGTCTCCACCGGT 120
DB	
QY	121 CCCCTGCCCAACCTCTGTGGCCCCACCTTCTTGGAGGCGCAACCCCGGGAGGAGCATTAG 180
DB	
QY	121 CCCCTGCCCAACCTCTGTGGCCCCACCTTCTTGGAGGCGCAACCCCGGGAGGAGCATTAG 180
DB	
QY	181 AAGGATTTTTCCGCGAGTTGCGAAGGAGCAAACTTGTGTGCAACTTGCCTCCCGGT 240
DB	
QY	181 AAGGATTTTTCCGCGAGTTGCGAAGGAGCAAACTTGTGTGCAACTTGCCTCCCGGT 240
DB	
QY	241 CGGGCGCTCTCCGCCCAACCTCTCAACATGCTTAGGGTCCGGGCGCCGGGCTGTCTGCT 300
DB	
QY	241 CGGGCGCTCTCCGCCCAACCTCTCAACATGCTTAGGGTCCGGGCGCCGGGCTGTCTGCT 300
DB	
QY	301 GCTGGCGTCTGTGCTCGGGGAGCGGTCCCTCCACGCGAGCTCTCAAGAGCAAGAG 360
DB	
QY	301 GCTGGCGTCTGTGCTCGGGGAGCGGTCCCTCCACGCGAGCTCTCAAGAGCAAGAG 360
DB	
QY	361 GCAGGCTCAGAAATGTTTCAGCCCCAGTCCCGGTGCTGTCAAGTCAAGCAAGCCCGG 420
DB	
QY	361 GCAGGCTCAGAAATGTTTCAGCCCCAGTCCCGGTGCTGTCAAGTCAAGCAAGCCCGG 420
DB	
QY	421 TTGTTATGCAATGGAATACTATCAGATAAATCAACAGTGGGAGCGGACCTACCTAGG 480
DB	
QY	421 TTGTTATGCAATGGAATACTATCAGATAAATCAACAGTGGGAGCGGACCTACCTAGG 480
DB	
QY	481 CAATGCGTTGTTTGTACTGTGTATGAGGAAGCCGAGGTTTAACTCGGAAGTAAACC 540
DB	
QY	481 CAATGCGTTGTTTGTACTGTGTATGAGGAAGCCGAGGTTTAACTCGGAAGTAAACC 540
DB	
QY	541 TGAAGCTGAAGAGACTTGCTTTGTGAAGATACTGTGGGAACACTTACCGAGTGGGTGACAC 600

1541 TGAAGCTGAAGAGACTGTTTGAACAAGTCACTGGGAACAATTACCGAGTGGGTGACAC 600
1601 TTATGAGCGCTCTAAAGACTCCATGATCTGGGACTGTACTCTGCATCGGGCTGGGCGAGG 660
1601 TTATGAGCGCTCTAAAGACTCCATGATCTGGGACTGTACTCTGCATCGGGCTGGGCGAGG 660
1661 GAGAAATAAGCTGTACCATCGCAACCGCTGCCATGAAGGGGTGAGTCTTACAGATTGG 720
1661 GAGAAATAAGCTGTACCATCGCAACCGCTGCCATGAAGGGGTGAGTCTTACAGATTGG 720
1721 TGACACCTGGAGGAGACCAATGAGACTGGTGTATCATGTTAGAGTGTGTGTCTTGG 780
1721 TGACACCTGGAGGAGACCAATGAGACTGGTGTATCATGTTAGAGTGTGTGTCTTGG 780
1781 TAATGGAAGAGAGAAATGGAACCTGCAAGCCATGAGTGAAGAGTGTGTTGATCATGCTGC 840
1781 TAATGGAAGAGAGAAATGGAACCTGCAAGCCATGAGTGAAGAGTGTGTTGATCATGCTGC 840
1841 TGGGACTTCTATGTGGTCCGAGAAAAGCTGGGAGAAAGCCCTACCAAGGCTGGATGATGGT 900
1841 TGGGACTTCTATGTGGTCCGAGAAAAGCTGGGAGAAAGCCCTACCAAGGCTGGATGATGGT 900
1901 AGATTGTACTTGGCTGGGAGAGCGAGCGGACGCATCATCTGCATCTTAGAAATAGATG 960
1901 AGATTGTACTTGGCTGGGAGAGCGAGCGGACGCATCATCTGCATCTTAGAAATAGATG 960
1961 CAACGATCAGGACACAAAGGACATCTTATGAAATGGAGACACCTGGAGCAAGAGATAA 1020
1961 CAACGATCAGGACACAAAGGACATCTTATGAAATGGAGACACCTGGAGCAAGAGATAA 1020
1021 TCGAGGAACCTGCTCCAGTGCATCTGCACAGCAACCGCGGAGAGTGGAGTGTGA 1080
1021 TCGAGGAACCTGCTCCAGTGCATCTGCACAGCAACCGCGGAGAGTGGAGTGTGA 1080
1081 GAGGACACCTCTGTGAGAGACCATCGAGCGGATCTGGCCCTTACCGGATGTTCTGTCG 1140
1081 GAGGACACCTCTGTGAGAGACCATCGAGCGGATCTGGCCCTTACCGGATGTTCTGTCG 1140
1141 AGCTGTTTACAAACCGGAGCTCACCCGAGCTCTCCCTATGGGCACTGTGTACAGA 1200
1141 AGCTGTTTACAAACCGGAGCTCACCCGAGCTCTCCCTATGGGCACTGTGTACAGA 1200
1201 CAGTGGTGTGCTTACTCTGTGGGATGCTAGTGGCTGAAGACACAAGGAAATGAAGAAAT 1260
1201 CAGTGGTGTGCTTACTCTGTGGGATGCTAGTGGCTGAAGACACAAGGAAATGAAGAAAT 1260
1261 GCTTTGCACTGCTGGGCAACCGAGTCACTGCTGCAAGACACAAGGAAATGAAGAAAT 1320
1261 GCTTTGCACTGCTGGGCAACCGAGTCACTGCTGCAAGACACAAGGAAATGAAGAAAT 1320
1321 CGGTGGCAACTCAAAATGGAGAGCCATGTGCTTACCAATTCACCTACAAATGGGAGAGT 1380
1321 CGGTGGCAACTCAAAATGGAGAGCCATGTGCTTACCAATTCACCTACAAATGGGAGAGT 1380
1381 CTACTCTCTGCAACCAAGAGGCGCAAGGACGGAATCTTTTGTGTGAGCAACAATTGGA 1440
1381 CTACTCTCTGCAACCAAGAGGCGCAAGGACGGAATCTTTTGTGTGAGCAACAATTGGA 1440
1441 TTATGAGCAGGACCAAGAAATACTCTTCTGACAGACCACTGTTTGTGTGAGACTCG 1500
1441 TTATGAGCAGGACCAAGAAATACTCTTCTGACAGACCACTGTTTGTGTGAGACTCG 1500
1501 AGGAGGAAATCCAAATGGTGTGCTGTCACCTTCCCTTCTTACAAACCAACAATTA 1560
1501 AGGAGGAAATCCAAATGGTGTGCTGTCACCTTCCCTTCTTACAAACCAACAATTA 1560
1561 CACTGATTGCACTTCTGAGGCAAGAGCAACATGAGTGGTGGGACCAACAGAA 1620
1561 CACTGATTGCACTTCTGAGGCAAGAGCAACATGAGTGGTGGGACCAACAGAA 1620
1621 CTATGATGCCACCAAGAGTGTGGTGTCTGCCCCATGGCTGCCACAGAGAAATCTGAC 1680

Db 1621 CTATGATGCCACAGAAAGTTTGGGTTCTGCCCATGGCTGCCACAGAGAAATCTGCAC 1680
QY 1681 AACCAATGAAGGGGTGATGTACCGCATTTGGAGATAGTGGGATAAGCAGCATGACATGG 1740
Db 1681 AACCAATGAAGGGGTGATGTACCGCATTTGGAGATAGTGGGATAAGCAGCATGACATGG 1740
QY 1741 TCACATGATCAGGTGACGCTGTGTTGGGAATGCTGTTGGGAATGGAACATGCAATTCGCTA 1800
Db 1741 TCACATGATCAGGTGACGCTGTGTTGGGAATGCTGTTGGGAATGGAACATGCAATTCGCTA 1800
QY 1801 CTGCGACGTTGAGATCAGTGCATTTGATGATGATCATCTTCAATGTGAAACGACAT 1860
Db 1801 CTGCGACGTTGAGATCAGTGCATTTGATGATGATCATCTTCAATGTGAAACGACAT 1860
QY 1861 CCACAAGCGTCATGAAGAGGGGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1920
Db 1861 CCACAAGCGTCATGAAGAGGGGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1920
QY 1921 CAGGTGAAGTGTGATCCGCTCGACCAATGCGAGGATTCAGAGCTGGGACGTTTATCA 1980
Db 1921 CAGGTGAAGTGTGATCCGCTCGACCAATGCGAGGATTCAGAGCTGGGACGTTTATCA 1980
QY 1981 AATTGAGATTCATGCGGAGAGTATGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2040
Db 1981 AATTGAGATTCATGCGGAGAGTATGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2040
QY 2041 CCGTGGCATTTGGGAGTGGCATTTGCCAACTTTACAGACTTCCAAAGCTCAAGTGTGCT 2100
Db 2041 CCGTGGCATTTGGGAGTGGCATTTGCCAACTTTACAGACTTCCAAAGCTCAAGTGTGCT 2100
QY 2101 TGTGGAAGTATTTATCACTGAGACTCCGAGTCCAGCCCACTCCACCCCATCCAGTGGAA 2160
Db 2101 TGTGGAAGTATTTATCACTGAGACTCCGAGTCCAGCCCACTCCACCCCATCCAGTGGAA 2160
QY 2161 TGCACCAAGCCATCTCACTTTCCAAAGTACATTTCCAGGTGGAGCTTAAATTTCTGT 2220
Db 2161 TGCACCAAGCCATCTCACTTTCCAAAGTACATTTCCAGGTGGAGCTTAAATTTCTGT 2220
QY 2221 AGGCGTTGGAAGGAAAGTACCATACAGGCGCACTTAAACTCTTACACCATCAAGGCGCT 2280
Db 2221 AGGCGTTGGAAGGAAAGTACCATACAGGCGCACTTAAACTCTTACACCATCAAGGCGCT 2280
QY 2281 GAAGCTGCTGTGTATACAGAGGCGCACTCATCAGCATCCAGCAGTACGCGCCACCAAGA 2340
Db 2281 GAAGCTGCTGTGTATACAGAGGCGCACTCATCAGCATCCAGCAGTACGCGCCACCAAGA 2340
QY 2341 AGTGACTCGCTTTGACTTCCACCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2400
Db 2341 AGTGACTCGCTTTGACTTCCACCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2400
QY 2401 GACGAGGAGAGCAGCTCCCTTTCTCTTGTGGCCACTTCTGAAATCTGTGACCGGAAAT 2460
Db 2401 GACGAGGAGAGCAGCTCCCTTTCTCTTGTGGCCACTTCTGAAATCTGTGACCGGAAAT 2460
QY 2461 CACGACGATAGCTTTGTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2520
Db 2461 CACGACGATAGCTTTGTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2520
QY 2521 GGTGGAATATCAGGTGAGTGGAGGAGGAGTGGCCACAGTACCTGGATCTTCCAAAGCAC 2580
Db 2521 GGTGGAATATCAGGTGAGTGGAGGAGGAGTGGCCACAGTACCTGGATCTTCCAAAGCAC 2580
QY 2581 AGCCACTTCTGTGAACATCCCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2640
Db 2581 AGCCACTTCTGTGAACATCCCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2640
QY 2641 TCAGATATCTGAGGATGGGAGCAGAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2700
Db 2641 TCAGATATCTGAGGATGGGAGCAGAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2700
QY 2701 TGATGCCCTCTGACCCGACTGTGGACCAAGTGTGACACCTCAATTTGTTGCTGCTG 2760
Db 2701 TGATGCCCTCTGACCCGACTGTGGACCAAGTGTGACACCTCAATTTGTTGCTGCTG 2760

QY 2761 GAGCAGCCCCAGGCTCCATCAGAGGTACAGATAGTCTATTGCGCATCAGTAGAAGG 2820
 DB 2761 GAGCAGCCCCAGGCTCCATCAGAGGTACAGATAGTCTATTGCGCATCAGTAGAAGG 2820
 QY 2821 TAGCAGCAGAGACTCAACCTTCCTGAAACTGCAAACTCCGTGACCCCTCAGTGACTTGCA 2880
 DB 2821 TAGCAGCAGAGACTCAACCTTCCTGAAACTGCAAACTCCGTGACCCCTCAGTGACTTGCA 2880
 QY 2881 ACCTGGTGTTCAGTATACATCATCTATCTGCTGTGGAGAAATCAAGAAAGTACACC 2940
 DB 2881 ACCTGGTGTTCAGTATACATCATCTATCTGCTGTGGAGAAATCAAGAAAGTACACC 2940
 QY 2941 TGTTGTCAATCAACAGAAACCACTGGCACCCACGCTCAGATACAGTCCCTCCTCCAG 3000
 DB 2941 TGTTGTCAATCAACAGAAACCACTGGCACCCACGCTCAGATACAGTCCCTCCTCCAG 3000
 QY 3001 GGACTGCAAGTTTGGAGGTGACAGAGTCAAGGTCAACCATCATGTGGACCGCCTGA 3060
 DB 3001 GGACTGCAAGTTTGGAGGTGACAGAGTCAAGGTCAACCATCATGTGGACCGCCTGA 3060
 QY 3061 GAGTCACTGACCGGCTACCGTGTGGATGTGATCCCGTCAACCTGGCGAGCACGG 3120
 DB 3061 GAGTCACTGACCGGCTACCGTGTGGATGTGATCCCGTCAACCTGGCGAGCACGG 3120
 QY 3121 GCAGAGGTGCCCATCAGCAGGAACCTTTGCAAGAGTCAACCGGCTGTCCCTGGGGT 3180
 DB 3121 GCAGAGGTGCCCATCAGCAGGAACCTTTGCAAGAGTCAACCGGCTGTCCCTGGGGT 3180
 QY 3181 CACCTATTACTTCAAGTCTTTGAGTGAAGTCAAGGTCAAGGTCAAGGTCAAGGTCAAGGT 3240
 DB 3181 CACCTATTACTTCAAGTCTTTGAGTGAAGTCAAGGTCAAGGTCAAGGTCAAGGTCAAGGT 3240
 QY 3241 TCAACAGAACCAAACTGGATGTCCCACTAACTCCAGTTTGTCAATGAATCATGATTC 3300
 DB 3241 TCAACAGAACCAAACTGGATGTCCCACTAACTCCAGTTTGTCAATGAATCATGATTC 3300
 QY 3301 TACTGTCTGTGATGAGTCACTCCAGTGTGATGATGATGATGATGATGATGATGATGAT 3360
 DB 3301 TACTGTCTGTGATGAGTCACTCCAGTGTGATGATGATGATGATGATGATGATGATGAT 3360
 QY 3361 GGGCTTACCAGAGAGAACAGCCAGGAGTACAAATGAGTCCCTGTCTCCAAAGTA 3420
 DB 3361 GGGCTTACCAGAGAGAACAGCCAGGAGTACAAATGAGTCCCTGTCTCCAAAGTA 3420
 QY 3421 CCCACTGAGGAATCTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3480
 DB 3421 CCCACTGAGGAATCTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3480
 QY 3481 CAACCAAGAGAGCCCCAAGCCACTGGAGTCTTTACCACTGAGGCTGGAGCTCTAT 3540
 DB 3481 CAACCAAGAGAGCCCCAAGCCACTGGAGTCTTTACCACTGAGGCTGGAGCTCTAT 3540
 QY 3541 TCCACCTTACACACCGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3600
 DB 3541 TCCACCTTACACACCGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3600
 QY 3601 AAGAAATGGTTTAAAGCTGGGTGTACGACCAAGCCAGGAGGAGGAGGAGGAGGAGGAGGAGG 3660
 DB 3601 AAGAAATGGTTTAAAGCTGGGTGTACGACCAAGCCAGGAGGAGGAGGAGGAGGAGGAGGAGG 3660
 QY 3661 GACTTCAGACTCAGGAAGCATCGTTGTGCTGGGTGACTCCAGGAGTGAATACGTTCTA 3720
 DB 3661 GACTTCAGACTCAGGAAGCATCGTTGTGCTGGGTGACTCCAGGAGTGAATACGTTCTA 3720
 QY 3721 CACCATCAAGTCTCAGAGATGAGCAGGAAGAGATGCGCAATTTGTAACAAAGTGGT 3780
 DB 3721 CACCATCAAGTCTCAGAGATGAGCAGGAAGAGATGCGCAATTTGTAACAAAGTGGT 3780
 QY 3781 GACACCATTTCTCCACCAACAACTTGCATCTGAGGAGCAACCCCTGACCTGAGTGCT 3840
 DB 3781 GACACCATTTCTCCACCAACAACTTGCATCTGAGGAGCAACCCCTGACCTGAGTGCT 3840

3841 CACAGTCTCTGGGAGAGGAGCACACCCAGACATTAAGTCTATAGATTAACACAAAC 3900
 DB 3841 CACAGTCTCTGGGAGAGGAGCACACCCAGACATTAAGTCTATAGATTAACACAAAC 3900
 QY 3901 CCCTACAAAGCCGACAGAGGAAATTTCTTTGGAAGAAGTGTCTCATGTCATGATCAGAGTTC 3960
 DB 3901 CCCTACAAAGCCGACAGAGGAAATTTCTTTGGAAGAAGTGTCTCATGTCATGATCAGAGTTC 3960
 QY 3961 CTGCACTTTTGTATAACCTCAGTCCCGGCTGGAGTACAAATGTCAATGTCACTGTCTCAA 4020
 DB 3961 CTGCACTTTTGTATAACCTCAGTCCCGGCTGGAGTACAAATGTCACTGTCTCAA 4020
 QY 4021 GGATGACAAAGAAAGTGTCTCTATCTCTGATACCATCATCCAGTGTCTCTCTCCAC 4080
 DB 4021 GGATGACAAAGAAAGTGTCTCTATCTCTGATACCATCATCCAGTGTCTCTCTCCAC 4080
 QY 4081 TGACCTTGCATTTACCAACATTTGGTCCAGACACCATGCTGTCACCTGGGCTCCACCCCC 4140
 DB 4081 TGACCTTGCATTTACCAACATTTGGTCCAGACACCATGCTGTCACCTGGGCTCCACCCCC 4140
 QY 4141 ATCCATTTGATTTAAACCACTTCTCTGCTGGTGTACTCACTGTGAAAAATGAGAAAGTGT 4200
 DB 4141 ATCCATTTGATTTAAACCACTTCTCTGCTGGTGTACTCACTGTGAAAAATGAGAAAGTGT 4200
 QY 4201 TGCAGAGTTGTCAATTTCTCTCAGACAAATGCTGCTTTTAAACAAATCTCTCTCTGG 4260
 DB 4201 TGCAGAGTTGTCAATTTCTCTCAGACAAATGCTGCTTTTAAACAAATCTCTCTCTGG 4260
 QY 4261 TACAGAAATGTAGTGAAGTCTCAGTGTCTTACGAAACAAATGAGAGACACCTCTTAG 4320
 DB 4261 TACAGAAATGTAGTGAAGTCTCAGTGTCTTACGAAACAAATGAGAGACACCTCTTAG 4320
 QY 4321 AGAAGACAGAAACAGGCTTTGATTTCCCACTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4380
 DB 4321 AGAAGACAGAAACAGGCTTTGATTTCCCACTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4380
 QY 4381 CAACTCTTTTACTGTGCACTGGATGCTCTCTGAGCCACCATCACTGCTGCTGCTGCTGCTGCT 4440
 DB 4381 CAACTCTTTTACTGTGCACTGGATGCTCTCTGAGCCACCATCACTGCTGCTGCTGCTGCTGCT 4440
 QY 4441 CCAATCTCCGAGCACTTCAGTGGGAGCTGAGAGATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4500
 DB 4441 CCAATCTCCGAGCACTTCAGTGGGAGCTGAGAGATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4500
 QY 4501 TTTCCATCACCCTCACCACCTCCTCCTGAGGACAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 4560
 DB 4501 TTTCCATCACCCTCACCACCTCCTCCTGAGGACAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 4560
 QY 4561 TAAATGCGAGAGGAAAGTCCCTTATTTGATTTGCGCAACAAATCAACAGTTTCTGATGTTCC 4620
 DB 4561 TAAATGCGAGAGGAAAGTCCCTTATTTGATTTGCGCAACAAATCAACAGTTTCTGATGTTCC 4620
 QY 4621 GAGGAGCTTGGAAAGTTGCTGCGACCCCTCCTGAGGACAGATGATGCTGCTGCTGCTGCTGCTGCTGCT 4680
 DB 4621 GAGGAGCTTGGAAAGTTGCTGCGACCCCTCCTGAGGACAGATGATGCTGCTGCTGCTGCTGCTGCTGCT 4680
 QY 4681 TGCTGTCAAGTGAAGTATTAAGGATCACTTACGAGGAAACAGGAGGAAATAGCCCTGT 4740
 DB 4681 TGCTGTCAAGTGAAGTATTAAGGATCACTTACGAGGAAACAGGAGGAAATAGCCCTGT 4740
 QY 4741 CAGGAGTTTCACTGCTGCTGCGGAGGAAAGTCTACGCTACCATCAGGCTTTAAACCTGG 4800
 DB 4741 CAGGAGTTTCACTGCTGCGGAGGAAAGTCTACGCTACCATCAGGCTTTAAACCTGG 4800
 QY 4801 AGTTGATTAATCAATCACTGTGTATGCTCTACTGCGCTGGAGACAGCCCGCAAGCAG 4860
 DB 4801 AGTTGATTAATCAATCACTGTGTATGCTCTACTGCGCTGGAGACAGCCCGCAAGCAG 4860
 QY 4861 CAGGCAATTTCCATTAATTCGACAGAAATTCGACAAACCATCCAGATCAAGTGAC 4920
 DB 4861 CAGGCAATTTCCATTAATTCGACAGAAATTCGACAAACCATCCAGATCAAGTGAC 4920
 QY 4921 CGATGTTTCCAGGACACAGCATTAAGTGTCAAGTGGCTGCTTCAAGTTTCCCTGTTACTGG 4980

4921	CGATGTTCCAGGACACAGCATAGTGTCAAGTGGTGCCTTCAAGTTCCTCCCTGTTACTGG	6001	CTTGGCCACCAACACCAATTCCTTGTCTGTATCATGGCAGCGCCACGTCGACGATTAC	6060
4981	TTACAGAGTAACACCACTCCCAAAATGGACACGAGCAACAAATAAATAAATAAATAAATAA	6061	CGGCTACATCATCAAGATATGAGAAGCCTGGTCTCTCTCCAGAGAGAGTGGTCCCTCGGCC	6120
4981	TTACAGAGTAACACCACTCCCAAAATGGACACGAGCAACAAATAAATAAATAAATAAATAA	6061	CGGCTACATCATCAAGATATGAGAAGCCTGGTCTCTCTCCAGAGAGAGTGGTCCCTCGGCC	6120
5041	TCCAGATCAAAACAGAAATGACTATTTGAAGGCTTGGAGCCACAGTGGAGTATGGTTAG	6121	CGGCTGTGTGTCACAGAGGCTTACTTATTACTGGCTTGAACCGGGAAACCGAATATACAAAT	6180
5041	TCCAGATCAAAACAGAAATGACTATTTGAAGGCTTGGAGCCACAGTGGAGTATGGTTAG	6121	CGGCTGTGTGTCACAGAGGCTTACTTATTACTGGCTTGAACCGGGAAACCGAATATACAAAT	6180
5101	TGCTCTATGCTCAGAAATCCAAAGCGGAGAGTCAAGTCTGCTTCCAGTCAAGTCAAAATTC	6181	TTATGTCTATGTCCTTGAAGATTAATCAGAGAGCAGAGCCCTGATTTGAAGGAAAGAAC	6240
5101	TGCTCTATGCTCAGAAATCCAAAGCGGAGAGTCAAGTCTGCTTCCAGTCAAGTCAAAATTC	6181	TTATGTCTATGTCCTTGAAGATTAATCAGAGAGCAGAGCCCTGATTTGAAGGAAAGAAC	6240
5161	CATTGATCGCCCTAAAGACTTGGCATTACATGATGTGGATGTCGATTCATCAAAATTC	6241	AGACGAGCTTCCCAACCTGGTAAACCTTCCACACCCCAATCTTCAATGACACGAGATCTT	6300
5161	CATTGATCGCCCTAAAGACTTGGCATTACATGATGTGGATGTCGATTCATCAAAATTC	6241	AGACGAGCTTCCCAACCTGGTAAACCTTCCACACCCCAATCTTCAATGACACGAGATCTT	6300
5221	TTGGGAAAGCCACAGGGGCAAGTTTCCAGGTACAGGGTGACCTACTCGAGCCCTGAGGA	6301	GGATGTTCTTCCACAGTTCAAAAGACCCCTTTCGTACCCACCCCTGGGTATGACACTGG	6360
5221	TTGGGAAAGCCACAGGGGCAAGTTTCCAGGTACAGGGTGACCTACTCGAGCCCTGAGGA	6301	GGATGTTCTTCCACAGTTCAAAAGACCCCTTTCGTACCCACCCCTGGGTATGACACTGG	6360
5281	TGGAAATCCATGAGCTATTCCTTGCACTGATGGTGAAGAGACACTGCGAGCTGCAAGG	6361	AAATGTATTCAGCTTCTGGCACCTTCTGGTCAGCAACCCAGTGTGGGCAACAATGAT	6420
5281	TGGAAATCCATGAGCTATTCCTTGCACTGATGGTGAAGAGACACTGCGAGCTGCAAGG	6361	AAATGTATTCAGCTTCTGGCACCTTCTGGTCAGCAACCCAGTGTGGGCAACAATGAT	6420
5341	CCTCAGACCGGGTCTGAGTACACAGTCAGTGTGGTTCCTTGCACGATGATATGGAGAG	6421	CTTTGAGGAACATGGTTTAAAGCGGACCAACCGGCCCAACCGGCCACCCCATTAAGGCA	6480
5341	CCTCAGACCGGGTCTGAGTACACAGTCAGTGTGGTTCCTTGCACGATGATATGGAGAG	6421	CTTTGAGGAACATGGTTTAAAGCGGACCAACCGGCCCAACCGGCCACCCCATTAAGGCA	6480
5401	CCAGCCCTGATTTGAACCCAGTCCACAGCTATTCCTGCAACAACTGACCTGAAAGTTCAAC	6481	TAGGCCAAGACCATACCCGCCAATGTAGGACAAAGAGTCTCTCAGACAAACCATCTC	6540
5401	CCAGCCCTGATTTGAACCCAGTCCACAGCTATTCCTGCAACAACTGACCTGAAAGTTCAAC	6481	TAGGCCAAGACCATACCCGCCAATGTAGGACAAAGAGTCTCTCAGACAAACCATCTC	6540
5461	TCAGGTCAACCCCAAGCCTGAGCGCCAGTGGACACACCCATGTTCACTGCTCACTGG	6541	ATGGGCCCATCTCCAGGACACTTCTGAGTACATCAATTTCAATCTATCTTGGTGGCACTGA	6600
5461	TCAGGTCAACCCCAAGCCTGAGCGCCAGTGGACACACCCATGTTCACTGCTCACTGG	6541	ATGGGCCCATCTCCAGGACACTTCTGAGTACATCAATTTCAATCTATCTTGGTGGCACTGA	6600
5521	ATATCGAGTGCGGGTGACCCCAAGGAGAGACCGGACCAATGAAGAAATCAACCTTGC	6601	TGAAGAACCTTACAGTTCAGGGTTCCTGGAACTTCTACAGTGGCACTCTGACAGGCT	6660
5521	ATATCGAGTGCGGGTGACCCCAAGGAGAGACCGGACCAATGAAGAAATCAACCTTGC	6601	TGAAGAACCTTACAGTTCAGGGTTCCTGGAACTTCTACAGTGGCACTCTGACAGGCT	6660
5581	TCCTGACAGCTCATCCGTTGTTATCAGGACTTATGTTGGCCCAACCAATGAAGTGAAG	6661	CACGAGAGTGCCACCTACACATCATAGTGAGGACCTGAAAGACCCAGCAGAGCATAA	6720
5581	TCCTGACAGCTCATCCGTTGTTATCAGGACTTATGTTGGCCCAACCAATGAAGTGAAG	6661	CACGAGAGTGCCACCTACACATCATAGTGAGGACCTGAAAGACCCAGCAGAGCATAA	6720
5641	TGCTCTATGCTTTAAGGACACTTTGACAGCAGACAGCTCAGGAGTGTGTCACCACTCT	6721	GGTTCGGGAGAGGTTGTTACCGTGGGCACTCTGTCAACGAGAGCTTGAACCAACCTAC	6780
5641	TGCTCTATGCTTTAAGGACACTTTGACAGCAGACAGCTCAGGAGTGTGTCACCACTCT	6721	GGTTCGGGAGAGGTTGTTACCGTGGGCACTCTGTCAACGAGAGCTTGAACCAACCTAC	6780
5701	GGAGAAATGTACGCCCAACCAAGAGGCTGTTGACAGATGCTACTGAGACCAATCAC	6781	GGATGACTCGTGTGTTGACCCCTTACACAGTTCCTTCCATTTCCATTTGAGATGAGTGGGA	6840
5701	GGAGAAATGTACGCCCAACCAAGAGGCTGTTGACAGATGCTACTGAGACCAATCAC	6781	GGATGACTCGTGTGTTGACCCCTTACACAGTTCCTTCCATTTCCATTTGAGATGAGTGGGA	6840
5761	CATTAGCTGGAGAACCAAGACTGAGACGATCACTGGCTTCCAGTGTGATCGGTTCCAGC	6841	ACGAATGTCTGAATCAGGCTTTAAACTGTGTGCGAGTCTTACAGTTCCTTGGAGTGGTCA	6900
5761	CATTAGCTGGAGAACCAAGACTGAGACGATCACTGGCTTCCAGTGTGATCGGTTCCAGC	6841	ACGAATGTCTGAATCAGGCTTTAAACTGTGTGCGAGTCTTACAGTTCCTTGGAGTGGTCA	6900
5821	CAATGGCCAGACTCCAAATCCAGAGAACCAATCAAGCCAGATGTCAGAGCTACCAATCAC	6901	TTTTCAGATGTGATTCATCTAGATGTGTCATGACATGCTGTAACATGATGAGTGGGA	6960
5821	CAATGGCCAGACTCCAAATCCAGAGAACCAATCAAGCCAGATGTCAGAGCTACCAATCAC	6901	TTTTCAGATGTGATTCATCTAGATGTGTCATGACATGCTGTAACATGATGAGTGGGA	6960
5881	TGCTCTTACAAACAGGCACTGACTACAGATCTACCTGTACACTTGAATGACATGCTCG	6961	GAAAGTGGGACCGTCAGGAGAAATGGCCAGATGATGAGTGCATGCTTGGGAAACGG	7020
5881	TGCTCTTACAAACAGGCACTGACTACAGATCTACCTGTACACTTGAATGACATGCTCG	6961	GAAAGTGGGACCGTCAGGAGAAATGGCCAGATGATGAGTGCATGCTTGGGAAACGG	7020
5941	GAGTCCCTGTGGTCAATCCAGCGCTTCCATGAGTCACTGCTGATGATGATGATGATGAT	7021	AAAAGGAGAAATCAAGTGTGACCTCATGAGGCAACGCTGTACGATGATGAGGAAACATA	7080
5941	GAGTCCCTGTGGTCAATCCAGCGCTTCCATGAGTCACTGCTGATGATGATGATGATGAT	7021	AAAAGGAGAAATCAAGTGTGACCTCATGAGGCAACGCTGTGTACGATGATGAGGAAACATA	7080
6001	CCTGGCCACCAACCAATTCCTTGTGTATCATGGCAGCGCCACGTCGACGATTAC	7081	CCAGTAGGAGAAACAGTGGCAGAGGAATATCTCGGTGCCATTTGCTCTGCAATGCTTT	7140
6001	CCTGGCCACCAACCAATTCCTTGTGTATCATGGCAGCGCCACGTCGACGATTAC	7081	CCAGTAGGAGAAACAGTGGCAGAGGAATATCTCGGTGCCATTTGCTCTGCAATGCTTT	7140

QY	3010	GTGTTGGAAGTGACAGACGTTGAAGGTCAACATCATGTGGACACCGCTGAGAGTGCAGT	3069	QY	4090	ATTACCAACATTTGTTCCAGACACCATCGTGTCACTGGGCTCCACCCCATCCATTGA	4149
DB	3104	GTGTTGGAAGTGACAGACGTTGAAGGTCAACATCATGTGGACACCGCTGAGAGTGCAGT	3163	DB	4184	ATTCACCAACATTTGTTCCAGACACCATCGTGTCACTGGGCTCCACCCCATCCATTGA	4243
QY	3070	GACCGGCTACCGTGTGATGTATCCCGGTCAACCTGCTGGGAGACCGGAGAGGCT	3129	QY	4150	TTTAAACCAACTTCCTGGTGCCTTACTCACCTGTGAAAAATGAGGAAGATGTTGAGAGTT	4209
DB	3164	GACCGGCTACCGTGTGATGTATCCCGGTCAACCTGCTGGGAGACCGGAGAGGCT	3223	DB	4244	TTTAAACCAACTTCCTGGTGCCTTACTCACCTGTGAAAAATGAGGAAGATGTTGAGAGTT	4303
QY	3130	GCCCATCAGCAGGAACACCTTTGAGAGAGTCAACCGGCTGTCCCTGGGCTCACCTATTA	3189	QY	4210	GTCAATTTCTCCTTCCAGCAATGAGTGGTCTTAAACAATCTCCCTGGCTGTGACAGATA	4269
DB	3224	GCCCATCAGCAGGAACACCTTTGAGAGAGTCAACCGGCTGTCCCTGGGCTCACCTATTA	3283	DB	4304	GTCAATTTCTCCTTCCAGCAATGAGTGGTCTTAAACAATCTCCCTGGCTGTGACAGATA	4363
QY	3190	CTTCAAAGTCTTTGAGTGAAGCCATGGGAGGAGAGCAGAGCTCTGACTCTCAACAGAC	3249	QY	4270	TGTAGTGAAGTGTCTCCAGTGTCTACGAAACAACATGAGAGCACACCTCTTAGAGGAAGACA	4329
DB	3284	CTTCAAAGTCTTTGAGTGAAGCCATGGGAGGAGAGCAGAGCTCTGACTCTCAACAGAC	3343	DB	4364	TGTAGTGAAGTGTCTCCAGTGTCTACGAAACAACATGAGAGCACACCTCTTAGAGGAAGACA	4423
QY	3250	AACCAAACTGGATGCTCCCATCAACCTCCAGTTTGTCAATGAAGTCAATCTACTGTCTCT	3309	QY	4330	GAAAACAGAGTCTTGATTTCCCAACTGGCAATGACATTTTCTGATATTTACTGCCAATCTTT	4389
DB	3344	AACCAAACTGGATGCTCCCATCAACCTCCAGTTTGTCAATGAAGTCAATCTACTGTCTCT	3403	DB	4424	GAAAACAGAGTCTTGATTTCCCAACTGGCAATGACATTTTCTGATATTTACTGCCAATCTTT	4483
QY	3310	GGTGAAGTGAATCTCACTCGGGGCCAGATTAACAGGATACCGACTGACCGTGGCCCTTAC	3369	QY	4390	TACTGTGACTGTGATTTGCTCTCGAGGCCACCATCACTGGCTACAGGATCCGGCATATCC	4449
DB	3404	GGTGAAGTGAATCTCACTCGGGGCCAGATTAACAGGATACCGACTGACCGTGGCCCTTAC	3463	DB	4484	TACTGTGACTGTGATTTGCTCTCGAGGCCACCATCACTGGCTACAGGATCCGGCATATCC	4543
QY	3370	CGAAGAGCAGCCAGCAGTCAATGTGGTCCCTCTGCTCCCAAGTACCCACTGAG	3429	QY	4450	CGAGCACTTCAGTGGAGACCTCGAGAAAGATCGGGTGCCTCACTCTCGGAATTCATCATC	4509
DB	3464	CGAAGAGCAGCCAGCAGTCAATGTGGTCCCTCTGCTCCCAAGTACCCACTGAG	3523	DB	4544	CGAGCACTTCAGTGGAGACCTCGAGAAAGATCGGGTGCCTCACTCTCGGAATTCATCATC	4603
QY	3430	GAATCTGAGCTGCATCTGAGTACACCGTATCCCTGCTGGGCCATAAAGGGCAACCAAGA	3489	QY	4510	CCTCACCACCTCACTCCAGGACACAGATGATGTGTGTCAGCATCGTTGCTTTAAATGGCAG	4569
DB	3524	GAATCTGAGCTGCATCTGAGTACACCGTATCCCTGCTGGGCCATAAAGGGCAACCAAGA	3583	DB	4604	CCTCACCACCTCACTCCAGGACACAGATGATGTGTGTCAGCATCGTTGCTTTAAATGGCAG	4663
QY	3490	GAGCCCCAAGCCATGAGTCTTTACCACACTGACGCTGGAGCTCTATTCCACCTTA	3549	QY	4570	AGAGGAAATCCCTTATTGATTGGCCAAACAATCAACAGTTTCTGATGTTCCGAGGAGCT	4629
DB	3584	GAGCCCCAAGCCATGAGTCTTTACCACACTGACGCTGGAGCTCTATTCCACCTTA	3643	DB	4664	AGAGGAAATCCCTTATTGATTGGCCAAACAATCAACAGTTTCTGATGTTCCGAGGAGCT	4723
QY	3550	CAACACCGAGTGAAGTGAAGCACCATTGTGATCACATGGACGCTGCTCCAAAGATTGG	3609	QY	4630	GGAGTGTGTGTGCGACCCCAACAGCTACTGATCAGCTGGGATGCTCTGCTGTGAC	4689
DB	3644	CAACACCGAGTGAAGTGAAGCACCATTGTGATCACATGGACGCTGCTCCAAAGATTGG	3703	DB	4724	GGAGTGTGTGTGCGACCCCAACAGCTACTGATCAGCTGGGATGCTCTGCTGTGAC	4783
QY	3610	TTTAAAGTGGGTGACACCAAGCCAGGAGGAGGACCAACGAGAGTGAATTCAGA	3669	QY	4690	AGTGAGATATTACAGGATCACTTACCGAGAAACAGGAGGAAATAGCCCTGTCCAGAGTT	4749
DB	3704	TTTAAAGTGGGTGACACCAAGCCAGGAGGAGGACCAACGAGAGTGAATTCAGA	3763	DB	4784	AGTGAGATATTACAGGATCACTTACCGAGAAACAGGAGGAAATAGCCCTGTCCAGAGTT	4843
QY	3670	CTCAGGAAGCATCGTTGTGCTCGGCTTGACTCCAGGAGTAGAATACGTTCAACCATCCA	3729	QY	4750	CAGTGTGCTGGAGCAAGTCTACAGCTACCATCAGCGCCTTAAACCTGGAGTTGATTA	4809
DB	3764	CTCAGGAAGCATCGTTGTGCTCGGCTTGACTCCAGGAGTAGAATACGTTCAACCATCCA	3823	DB	4844	CAGTGTGCTGGAGCAAGTCTACAGCTACCATCAGCGCCTTAAACCTGGAGTTGATTA	4903
QY	3730	AGTCTGAGAGTGAAGCAGGAAGAGATGCGCCATTTGTAACAAGTGTGACACCATTT	3789	QY	4810	TACATCACTGTGTATGCTGTCACTGGCCGTGGAGACAGCCCGCAAGCAGCAAGCCAAAT	4869
DB	3824	AGTCTGAGAGTGAAGCAGGAAGAGATGCGCCATTTGTAACAAGTGTGACACCATTT	3883	DB	4904	TACATCACTGTGTATGCTGTCACTGGCCGTGGAGACAGCCCGCAAGCAGCAAGCCAAAT	4963
QY	3790	GTCTCCACCAAACTTGCATCTGGAGCAAAACCTGACACTGGAGTGTCAAGTCTC	3849	QY	4870	TTCCATTAATTACCGAACAGAAATTGAACAACTCCAGATGCAAGTGAACCGATGTTCA	4929
DB	3884	GTCTCCACCAAACTTGCATCTGGAGCAAAACCTGACACTGGAGTGTCAAGTCTC	3943	DB	4964	TTCCATTAATTACCGAACAGAAATTGAACAACTCCAGATGCAAGTGAACCGATGTTCA	5023
QY	3850	CTGGAGAGGAGCACCAACCCAGACATTAAGTGTATAGAAATTACCAACCCCTACAAA	3909	QY	4930	GGACACAGCATTTAGTGTCAAGTGGCTGCTTCAAGTTCCTCTGTACTGGTTACAGAGT	4989
DB	3944	CTGGAGAGGAGCACCAACCCAGACATTAAGTGTATAGAAATTACCAACCCCTACAAA	4003	DB	5024	GGACACAGCATTTAGTGTCAAGTGGCTGCTTCAAGTTCCTCTGTACTGGTTACAGAGT	5083
QY	3910	CGGCCAGCAGGGAATTTCTTGGAGAGAGTGTGCATGCTGATCAGAGCTCTCTGACCTTT	3969	QY	4990	AACCAACCTCCCAAAATGGACAGGACCAACAACTAATACTGAGGTCAGATCA	5049
DB	4004	CGGCCAGCAGGGAATTTCTTGGAGAGAGTGTGCATGCTGATCAGAGCTCTCTGACCTTT	4063	DB	5084	AACCAACCTCCCAAAATGGACAGGACCAACAACTAATACTGAGGTCAGATCA	5143
QY	3970	TGATAACCTGAGTCCCGGCTGGAGTACAATGTAGTGTGTTTACACTGTCAAGGATGACAA	4029	QY	5050	AACAGAAATGACTTTTGAAGGCTTGGAGCCCAAGTGGAGTATGTTGTTGTTCTATGC	5109
DB	4064	TGATAACCTGAGTCCCGGCTGGAGTACAATGTAGTGTGTTTACACTGTCAAGGATGACAA	4123	DB	5144	AACAGAAATGACTTTTGAAGGCTTGGAGCCCAAGTGGAGTATGTTGTTGTTCTATGC	5203
QY	4030	GGAAAGTGTCCCTATCTGTGATACCATCATCCAGCTGCTCTCTCCACTGACCTGCG	4089	QY	5110	TCAGAAATCCAAGCGGAGAGAGTCAAGCTCTGTTGAGTGTGAGTAAACCAATGATCG	5169
DB	4124	GGAAAGTGTCCCTATCTGTGATACCATCATCCAGCTGCTCTCTCCACTGACCTGCG	4183	DB	5204	TCAGAAATCCAAGCGGAGAGAGTCAAGCTCTGTTGAGTGTGAGTAAACCAATGATCG	5263
				QY	5170	CCCTAAAGGACTGGCAATTCATGATGTGATGCTGATTCATCAAAATTTGCTTGGAAAG	5229

b 5264 CCTAAGAGCTGGCATTACATGATGTGATGTGATTCATCAAAATGCTTGGAAAG 5323
 y 5230 CCCACAGGGGCAAGTTTCAGGTACAGGTGACCTACCTGAGCCCTGAGGATGGAATCCA 5289
 b 5324 CCCACAGGGGCAAGTTTCAGGTACAGGTGACCTACCTGAGCCCTGAGGATGGAATCCA 5383
 y 5290 TGAGCTATTCCCTGCACTGATGAGTGAAGACACCTGACAGAGCTGCAAGGCTCAGACC 5349
 b 5384 TGAGCTATTCCCTGCACTGATGAGTGAAGACACCTGACAGAGCTGCAAGGCTCAGACC 5443
 y 5350 GGGTTCAGTACACAGTCAGTGTGAGTGGTTCCTGACAGATGATGAGAGCCAGCCCT 5409
 b 5444 GGGTTCAGTACACAGTCAGTGTGAGTGGTTCCTGACAGATGATGAGAGCCAGCCCT 5503
 y 5410 GATTGGAACCCAGTCACAGCTATTCTCTGACCACTGACCTGGAAGTTTCACTCAGGTAC 5469
 b 5504 GATTGGAACCCAGTCACAGCTATTCTCTGACCACTGACCTGGAAGTTTCACTCAGGTAC 5563
 y 5470 ACCACAAGCCTGAGCGCCAGTGGACACCAACCAATGTTTCACTGACCTGAGATTCGAGT 5529
 b 5564 ACCACAAGCCTGAGCGCCAGTGGACACCAACCAATGTTTCACTGACCTGAGATTCGAGT 5623
 y 5530 GGGGTGACCCCAAGGAGACCGGACCAATGAAGAAATCAACCTTGCTCTGACAG 5589
 b 5624 GGGGTGACCCCAAGGAGACCGGACCAATGAAGAAATCAACCTTGCTCTGACAG 5683
 y 5590 CTCATCCGTGGTGTATCAGGACTTATGTTGGGCCACCAAAATGTAAGTGAAGTGTCTATGC 5649
 b 5684 CTCATCCGTGGTGTATCAGGACTTATGTTGGGCCACCAAAATGTAAGTGAAGTGTCTATGC 5743
 y 5650 TCTTAAGACACCTTTCAGACGACAGCTCAGGAGTGTGCAACCTCTCGGAATGT 5709
 b 5744 TCTTAAGACACCTTTCAGACGACAGCTCAGGAGTGTGCAACCTCTCGGAATGT 5803
 y 5710 CAGCCACCAAGAGGGCTCGTGTGACAGATGCTACTGAGACCAACCATCACCATTAGCTG 5769
 b 5804 CAGCCACCAAGAGGGCTCGTGTGACAGATGCTACTGAGACCAACCATCACCATTAGCTG 5863
 y 5770 GAGAACCAAGACTGAGACGATCAGTGGCTCCAGTTGATGGCTTCAGCCAAATGGCCA 5829
 b 5864 GAGAACCAAGACTGAGACGATCAGTGGCTCCAGTTGATGGCTTCAGCCAAATGGCCA 5923
 y 5830 GACTCCAATCCAGAGAACCATCAAGCAGATGTCAGAGCTACACCATCACTGGCTTACA 5889
 b 5924 GACTCCAATCCAGAGAACCATCAAGCAGATGTCAGAGCTACACCATCACTGGCTTACA 5983
 y 5890 ACCAGGACTGACTACAGATCTACCTGACAGCTGACCTTGAATGACATGCTCGAGCTCCCC 5949
 b 5984 ACCAGGACTGACTACAGATCTACCTGACAGCTGACCTTGAATGACATGCTCGAGCTCCCC 6043
 y 5950 TGTGGTCACTCGAGCCTCCACTGCCATTGATGACCATCCAACTCGCTTCTCGGCCAC 6009
 b 6044 TGTGGTCACTCGAGCCTCCACTGCCATTGATGACCATCCAACTCGCTTCTCGGCCAC 6103
 y 6010 CACACCAATTCCTTGTGTATGATGGACGCGCCACAGTGCAGGATTAACGGCTACAT 6069
 b 6104 CACACCAATTCCTTGTGTATGATGGACGCGCCACAGTGCAGGATTAACGGCTACAT 6163
 y 6070 CATCAAGTATGAGAGCTGGGTCTCTCCAGAGAGTGGTCCCTCGGCCCGCCCTGG 6129
 b 6164 CATCAAGTATGAGAGCTGGGTCTCTCCAGAGAGTGGTCCCTCGGCCCGCCCTGG 6223
 y 6130 TGTACAGAGGCTACTATTACTGGCTGGAAACCGGAAACCGAATATACAAATTATGTCTAT 6189
 b 6224 TGTACAGAGGCTACTATTACTGGCTGGAAACCGGAAACCGAATATACAAATTATGTCTAT 6283
 y 6190 TGGCCTGAAGAAATATCAGAGAGCGAGCCCTGATTGGAAGGAAAAGACAGAGAGCT 6249
 b 6284 TGGCCTGAAGAAATATCAGAGAGCGAGCCCTGATTGGAAGGAAAAGACAGAGAGCT 6343
 y 6250 TCCCAACTGGTAACCCCTCCACACCCCAATCTTCATGGACCAAGATCTTGGATTTCC 6309

Db 6344 TCCCAACTGGTAACCCCTCCACACCCCAATCTTCATGACCAAGAGATCTTGGATGTTCC 6403
 Qy 6310 TTCCACAGTTCAAAGACCCCTTTCTGTCACCCACCTCGGTATGACACTGGAAATGGTAT 6369
 Db 6404 TTCCACAGTTCAAAGACCCCTTTCTGTCACCCACCTCGGTATGACACTGGAAATGGTAT 6463
 Qy 6370 TCAGCTTCTCTGGCACTTCTGTGTCAGCAACCCAGTGTGGGCAACAAATGATCTTTGAGGA 6429
 Db 6464 TCAGCTTCTCTGGCACTTCTGTGTCAGCAACCCAGTGTGGGCAACAAATGATCTTTGAGGA 6523
 Qy 6430 ACATGGTTTTAGCGGACACACCGCCCAACCGGCCACCCCATAGGCATAGGCCAAG 6489
 Db 6524 ACATGGTTTTAGCGGACACACCGCCCAACCGGCCACCCCATAGGCATAGGCCAAG 6583
 Qy 6490 ACCATACCCCGCAATGTAGGACAAAGAGCTCTCTCTGACAAACCATCTCATTGGGCCCC 6549
 Db 6584 ACCATACCCCGCAATGTAGGACAAAGAGCTCTCTCTGACAAACCATCTCATTGGGCCCC 6643
 Qy 6550 ATTCCAGGACACTTCTGAGTACATCATTTTCATGTCATCTGTTGGCACTGATGAAGAACC 6609
 Db 6644 ATTCCAGGACACTTCTGAGTACATCATTTTCATGTCATCTGTTGGCACTGATGAAGAACC 6703
 Qy 6610 CTTACAGTTTCAAGGGTTCCTGGAACTTCTACAGTGGCACTCTGACAGGCTCTCACAGAGG 6669
 Db 6704 CTTACAGTTTCAAGGGTTCCTGGAACTTCTACAGTGGCACTCTGACAGGCTCTCACAGAGG 6763
 Qy 6670 TGCACCTACAAATCATNTAGTGGAGCACTGAAAGACCAAGAGGAGGATAGGTTGGGA 6729
 Db 6764 TGCACCTACAAATCATNTAGTGGAGCACTGAAAGACCAAGAGGAGGATAGGTTGGGA 6823
 Qy 6730 AGAGTTTGTACCGTGGCCAACTCTGTCAACGAAGGCTTGAACCAACCTACCGATGACTC 6789
 Db 6824 AGAGTTTGTACCGTGGCCAACTCTGTCAACGAAGGCTTGAACCAACCTACCGATGACTC 6883
 Qy 6790 GTGTTTTGACCCCTACACAGTGTCCCATNTAGTGGAGTGGAGATGAGTGGAAACGATGTC 6849
 Db 6884 GTGTTTTGACCCCTACACAGTGTCCCATNTAGTGGAGTGGAGATGAGTGGAAACGATGTC 6943
 Qy 6850 TGAATCAGGCTTTTAAACTGTTGTCAGTCTTGAAGTGGTCAATTTCAAGATG 6909
 Db 6944 TGAATCAGGCTTTTAAACTGTTGTCAGTCTTGAAGTGGTCAATTTCAAGATG 7003
 Qy 6910 TGAATCAGTGTGAGTGTGTCATGACAAATGTTGAACTCAAGATTTGGAGAGTGGGA 6969
 Db 7004 TGAATCAGTGTGAGTGTGTCATGACAAATGTTGAACTCAAGATTTGGAGAGTGGGA 7063
 Qy 6970 CCGTCAGGAGAAATGGCCAGATGATGAGTGCATGCTTGGGAAACGAAAGAGGA 7029
 Db 7064 CCGTCAGGAGAAATGGCCAGATGATGAGTGCATGCTTGGGAAACGAAAGAGGA 7123
 Qy 7030 ATTCAAGTGTGACCCCTCATGAGCAACGTTTACGATGATGGGAAACATACCACTAGG 7089
 Db 7124 ATTCAAGTGTGACCCCTCATGAGCAACGTTTACGATGATGGGAAACATACCACTAGG 7183
 Qy 7090 AGAAGTGTGAGAGGAAATATCTCGGTGCAATTTGCTCTGACATGCTTGGAGGCCA 7149
 Db 7184 AGAAGTGTGAGAGGAAATATCTCGGTGCAATTTGCTCTGACATGCTTGGAGGCCA 7243
 Qy 7150 CCGGGCTGGCGCTGTGACAACTGCGCAGACCTGGGGGTGAACCCAGTCCCGAAGGCAC 7209
 Db 7244 CCGGGCTGGCGCTGTGACAACTGCGCAGACCTGGGGGTGAACCCAGTCCCGAAGGCAC 7303
 Qy 7210 TACTGGCCGTCCTTACAAACAGATATCTCAGAGATACCATCAGAGACCAACATATGT 7269
 Db 7304 TACTGGCCGTCCTTACAAACAGATATCTCAGAGATACCATCAGAGACCAACATATGT 7363
 Qy 7270 TAATTTGCCAATTTGAGTGTCTTTCATGCTTTAGATGTAAGGCTGACAGAGAGATTTCCCG 7329
 Db 7364 TAATTTGCCAATTTGAGTGTCTTTCATGCTTTAGATGTAAGGCTGACAGAGAGATTTCCCG 7423
 Qy 7330 AGAGTAAATCATCTTTTCCAATCCAGAGAACAGATGCTCTCTCTGCAAGATCCATCTA 7389
 Db 7424 AGAGTAAATCATCTTTTCCAATCCAGAGAACAGATGCTCTCTCTGCAAGATCCATCTA 7483

1383 ACTCCTGCACACAGAGGGGACAGGACGGACATCTTTTGTGTCAGACACAACTTCGAATT 1442
 1382 ACTCCTGCACACAGAGGGGACAGGACGGACATCTTTTGTGTCAGACACAACTTCGAATT 1441
 1443 ATGACAGGACAGAAAATACCTTTCTGTCAGACACACACTGTTTGGTTTCAGACTCGAG 1502
 1442 ATGACAGGACAGAAAATACCTTTCTGTCAGACACACACTGTTTGGTTTCAGACTCGAG 1501
 1503 GAGGAAATTCGAATGGTGGCTTTGTCGCTTCCCTTCTCTATACAAACACAACTTACA 1562
 1502 GAGGAAATTCGAATGGTGGCTTTGTCGCTTCCCTTCTCTATACAAACACAACTTACA 1561
 1563 CTGATTGCACTTCAGGACAGACAGACAAATGAGTGGTGGTGGACACAGAACT 1622
 1562 CTGATTGCACTTCAGGACAGACAGACAAATGAGTGGTGGTGGACACAGAACT 1621
 1623 ATGATGCCACACAGAAATTTGGTGTCTGCCCATGGCTGCCACAGAGAAATCTGCACAA 1682
 1622 ATGATGCCACAGAAATTTGGTGTCTGCCCATGGCTGCCACAGAGAAATCTGCACAA 1681
 1683 CCAATGAGGGTTCATGTACCCATTTGGGATGAGTGGGATTAAGCAGCATGACATGGGTC 1742
 1682 CCAATGAGGGTTCATGTACCCATTTGGGATGAGTGGGATTAAGCAGCATGACATGGGTC 1741
 1743 ACATGATGAGTGCACGTCTGTGGGATGGTGGTGGGATGGACATGGCTTACT 1802
 1742 ACATGATGAGTGCACGTCTGTGGGATGGTGGTGGGATGGACATGGCTTACT 1801
 1803 CCGAGCTTCGAGATCAGTGCATTTGATGACATCACTTACAAATGTAACGACATTC 1862
 1802 CCGAGCTTCGAGATCAGTGCATTTGATGACATCACTTACAAATGTAACGACATTC 1861
 1863 ACAGCGTTCATGAGAGGGGACATCTGCACTGTACATCTTCTGGTTCAGGGTTCGGGCA 1922
 1862 ACAGCGTTCATGAGAGGGGACATCTGCACTGTACATCTTCTGGTTCAGGGTTCGGGCA 1921
 1923 GGTGGAAGTGTGATCCGTCGACCAATGCGAGATTCAGAGCTGGGACGTTTATCAAA 1982
 1922 GGTGGAAGTGTGATCCGTCGACCAATGCGAGATTCAGAGCTGGGACGTTTATCAAA 1981
 1983 TTGGGATTCATGGGAGAGTATGTCATGGTGTGATGACATCCAGTGTCTATGGC 2042
 1982 TTGGGATTCATGGGAGAGTATGTCATGGTGTGATGACATCCAGTGTCTATGGC 2041
 2043 GTGSCATTTGGGAGTGGCATTCGCAACCTTTACAGACCTATCCAAAGCTCAAGTGTCTG 2102
 2042 GTGSCATTTGGGAGTGGCATTCGCAACCTTTACAGACCTATCCAAAGCTCAAGTGTCTG 2101
 2103 TCGAAGTATTTATCACTGAGACTCCGAGTCCAGCCAACTCCACCCCATCCAGTGGAAATG 2162
 2102 TCGAAGTATTTATCACTGAGACTCCGAGTCCAGCCAACTCCACCCCATCCAGTGGAAATG 2161
 2163 CACACAGCCATCTACATTTCCAAATGATCTCAGTGGGAGACCTAAATTTCTGTAG 2222
 2162 CACACAGCCATCTACATTTCCAAATGATCTCAGTGGGAGACCTAAATTTCTGTAG 2221
 2223 GCGTTTGGAGGAGCTACATACAGGACCTTAACTCTTACACCTACCAAGGCTTGA 2282
 2222 GCGTTTGGAGGAGCTACATACAGGACCTTAACTCTTACACCTACCAAGGCTTGA 2281
 2283 AGCTGGTGTGATACAGGGGCGAGTCACTAGCATCCAGCATCCAGGACCTCCAAAGAG 2342
 2282 AGCTGGTGTGATACAGGGGCGAGTCACTAGCATCCAGCATCCAGGACCTCCAAAGAG 2341
 2343 TGACTCGCTTTGATCTACACACAGGACGACACCTGTGACGACGACACCGTGA 2402
 2342 TGACTCGCTTTGATCTACACACAGGACGACACCTGTGACGACGACACCGTGA 2401
 2403 CAGGAGAGCAGCTCCCTTTTCTCTGTCGCTTCTGAAATCTGTGACCGAAATCA 2462
 2402 CAGGAGAGCAGCTCCCTTTTCTCTGTCGCTTCTGAAATCTGTGACCGAAATCA 2461
 2463 CAGCCAGTAGCTTTGTGGTCTCTCTGGTCTCAGCTTCGACACCGGTGTCGGGATCCGGG 2522

2462 CAGCCAGTAGCTTTGTGGTCTCTGGTCTCAGCTTCGACACCGGTGTCGGGATTCGGGG 2521
 2523 TGAATATCAGCTGAGTGAAGGAGGAGATGACCCACAGTACCTGATCTTCAAGACAG 2582
 2522 TGAATATCAGCTGAGTGAAGGAGGAGATGACCCACAGTACCTGATCTTCAAGACAG 2581
 2583 CACTTCTGTGAACATCCCTGACCTGCTTCTTGGCGGAAATACATTTGTAATGTCATC 2642
 2582 CACTTCTGTGAACATCCCTGACCTGCTTCTTGGCGGAAATACATTTGTAATGTCATC 2641
 2643 AGATATCTCAGATGGGAGCAGAGTTTGTATCTGTCTACTTTCACAAAACACAGGCGCTG 2702
 2642 AGATATCTCAGATGGGAGCAGAGTTTGTATCTGTCTACTTTCACAAAACACAGGCGCTG 2701
 2703 ATGCCCTCTGTGACCCGACTGTGGACCAAGTTGATGACACTCAATTTGTTGGTGGTGA 2762
 2702 ATGCCCTCTGTGACCCGACTGTGGACCAAGTTGATGACACTCAATTTGTTGGTGGTGA 2761
 2763 GCAGACCCAGGCTCCCATCAAGGATCAGAAATAGTCTATTTCGCCATCAGTAGAAGGTA 2822
 2762 GCAGACCCAGGCTCCCATCAAGGATCAGAAATAGTCTATTTCGCCATCAGTAGAAGGTA 2821
 2823 GCAGACAGAACTCAACCTTCTGAAACTGCAAACTCCCTGACCTCAGTGTGCTCAAC 2882
 2822 GCAGACAGAACTCAACCTTCTGAAACTGCAAACTCCCTGACCTCAGTGTGCTCAAC 2881
 2883 CTGGTGTTCAGTATAATCATCTATCTGTGTGGAAGAAATCAAGAAAGTACACCTG 2942
 2882 CTGGTGTTCAGTATAATCATCTATCTGTGTGGAAGAAATCAAGAAAGTACACCTG 2941
 2943 TTGTCATTCAACAGAAACCACTGGCACCCCACTCAGATACAGTGGCTCTCCAGGG 3002
 2942 TTGTCATTCAACAGAAACCACTGGCACCCCACTCAGATACAGTGGCTCTCCAGGG 3001
 3003 ACCTGCACTTTGTGGAAGTACAGAGCTGGAAGTCACTATGTGGAACACCGCTGAGA 3062
 3002 ACCTGCACTTTGTGGAAGTACAGAGCTGGAAGTCACTATGTGGAACACCGCTGAGA 3061
 3063 GTGCACTGACCGGCTACCGTGTGGATGTGATCCCTGCAACCTGCTGGGAGCAGCGGC 3122
 3062 GTGCACTGACCGGCTACCGTGTGGATGTGATCCCTGCAACCTGCTGGGAGCAGCGGC 3121
 3123 AGAGCTGCCCATCAGCAGAAACACTTTGCAAGTCAACCGGCTGCTCCCTGGGCTCA 3182
 3122 AGAGCTGCCCATCAGCAGAAACACTTTGCAAGTCAACCGGCTGCTCCCTGGGCTCA 3181
 3183 CCTATTACTTCAAAAGTCTTTGCAAGTCAACCGGCTGCAACCGCTCTGACTGCTC 3242
 3182 CCTATTACTTCAAAAGTCTTTGCAAGTCAACCGGCTGCAACCGCTCTGACTGCTC 3241
 3243 AACAGACAAACAACTGGATGCTCCCACTAACTCCAGTTTGTCAATGAACTGATCTTA 3302
 3242 AACAGACAAACAACTGGATGCTCCCACTAACTCCAGTTTGTCAATGAACTGATCTTA 3301
 3303 CTGTCTCTGTGAGATGAGCTCCACCTCGGCGCCAGATAACAGGATACCGACTACCGTGG 3362
 3302 CTGTCTCTGTGAGATGAGCTCCACCTCGGCGCCAGATAACAGGATACCGACTACCGTGG 3361
 3363 GCCTTACCCGAGGAGCAGCCAGGAGTACATGTTGGTGGCTCTGCTCCAGTACC 3422
 3362 GCCTTACCCGAGGAGCAGCCAGGAGTACATGTTGGTGGCTCTGCTCCAGTACC 3421
 3423 CACTGAGGATCTCGACCTGCTCATCTGATGACACCGGTATCCCTCGTGGCCATAAGGGCA 3482
 3422 CACTGAGGATCTCGACCTGCTCATCTGATGACACCGGTATCCCTCGTGGCCATAAGGGCA 3481
 3483 ACCAAGAGAGCCCAAGGACCTGAGTCTTTACCACTGTCAGCTGGAGCTCTATTTC 3542
 3482 ACCAAGAGAGCCCAAGGACCTGAGTCTTTACCACTGTCAGCTGGAGCTCTATTTC 3541
 3543 CACTTACCAACCGAGGAGTGAAGTGAAGACCACTGATGATCAATGACGCTGCTCCAA 3602

3542 CACCTTAAACACCGAGGTGACTGAGACACCAATCGTGATCAATGGAGCGCTGCTCCAA 3601
 3603 GAATTTGGTTTTAAAGCTGGGTGTACGACCAAGCCAGGAGGAGAGGCCACCAACGAGAAGTGA 3662
 3602 GAATTTGGTTTTAAAGCTGGGTGTACGACCAAGCCAGGAGGAGAGGCCACCAACGAGAAGTGA 3661
 3663 CTTTCAGACTCAGGAAGCATCGTTGTGTCCGGCTTGACTCCAGGAGTAGAATAGCTCTACA 3722
 3662 CTTTCAGACTCAGGAAGCATCGTTGTGTCCGGCTTGACTCCAGGAGTAGAATAGCTCTACA 3721
 3723 CCATCCAAAGTCCTGAGAGATGGACAGAGAAAGAGATGCGCAATTGTAACAAAGTGGTGA 3782
 3722 CCATCCAAAGTCCTGAGAGATGGACAGAGAAAGAGATGCGCAATTGTAACAAAGTGGTGA 3781
 3783 CACCATTTCTCCACCAACAACTTGTCATCTGGAGGCAACCCCTGACCTGGAGTGGCTCA 3842
 3782 CACCATTTCTCCACCAACAACTTGTCATCTGGAGGCAACCCCTGACCTGGAGTGGCTCA 3841
 3843 CAGTCTCTCTGGAGAGGAGCACCACCCAGACATTTACTGGTTATAGAAATTTACCACAACCC 3902
 3842 CAGTCTCTCTGGAGAGGAGCACCACCCAGACATTTACTGGTTATAGAAATTTACCACAACCC 3901
 3903 CTACAAACGGCAGCAGGAAATTTCTTTGGAAGAGTGGTCCATGCTGATCAGAGTCCCT 3962
 3902 CTACAAACGGCAGCAGGAAATTTCTTTGGAAGAGTGGTCCATGCTGATCAGAGTCCCT 3961
 3963 GCATTTTGTATTAACCTGAGTCCCGGCTGGAGTACAAATGTCAAGTGTATTAACCTGTCAAGG 4022
 3962 GCATTTTGTATTAACCTGAGTCCCGGCTGGAGTACAAATGTCAAGTGTATTAACCTGTCAAGG 4021
 4023 ATGACAGGAAAGTGTCCCTATCTCTGATACCATCATCCAGCTGTTCCTCTCCACTG 4082
 4022 ATGACAGGAAAGTGTCCCTATCTCTGATACCATCATCCAGCTGTTCCTCTCCACTG 4081
 4083 ACCTGCGATTACCAACATTTGCTCCAGACACCATGCTGTCACCTGGGCTCCACCCCAT 4142
 4082 ACCTGCGATTACCAACATTTGCTCCAGACACCATGCTGTCACCTGGGCTCCACCCCAT 4141
 4143 CCATTTGATTAACCACTTCTGCTGGTGTACTCACTGTGAAATAATGAGGAAGATGTTG 4202
 4142 CCATTTGATTAACCACTTCTGCTGGTGTACTCACTGTGAAATAATGAGGAAGATGTTG 4201
 4203 CAGAGTGTCAATTTCTCTTCCAGCAATGAGTGTCTTAAACAAATCTCTGCTGGTA 4262
 4202 CAGAGTGTCAATTTCTCTTCCAGCAATGAGTGTCTTAAACAAATCTCTGCTGGTA 4261
 4263 CAGAAATATGATGAGTGTCTTCCAGTGTCTACGAAACCAATGAGGACACCTCTTAGAG 4322
 4262 CAGAAATATGATGAGTGTCTTCCAGTGTCTACGAAACCAATGAGGACACCTCTTAGAG 4321
 4323 GAAGACAGAAACAGGCTTGATTTCCCACTGGCATTTCTTCTGATATTACTGCCA 4382
 4322 GAAGACAGAAACAGGCTTGATTTCCCACTGGCATTTCTTCTGATATTACTGCCA 4381
 4383 ACTCTTTTACTGTGCACTGGATTGCTCTCTGAGCCACCATCTGCTACAGGATCCGCC 4442
 4382 ACTCTTTTACTGTGCACTGGATTGCTCTCTGAGCCACCATCTGCTACAGGATCCGCC 4441
 4443 ATCATCCGAGACATTCAGTGGGAGACCTCGAGAAGATCGGGTCCCACTCTCGGAATT 4502
 4442 ATCATCCGAGACATTCAGTGGGAGACCTCGAGAAGATCGGGTCCCACTCTCGGAATT 4501
 4503 CCATCACCTCACCACCTCACTCCAGGACAGAGTATGTTGTCAGCATCGTTGCTCTTA 4562
 4502 CCATCACCTCACCACCTCACTCCAGGACAGAGTATGTTGTCAGCATCGTTGCTCTTA 4561
 4563 ATGGCAGAGAGAAAGTCCCTTTATGATGGCCAAACAAATCAACAGTTTCTGATGTTCCGA 4622
 4562 ATGGCAGAGAGAAAGTCCCTTTATGATGGCCAAACAAATCAACAGTTTCTGATGTTCCGA 4621
 4623 GGGACCTGGAAGTGTGCTGCGACCCCAACAGCTTACTGATCAGCTGGAGTGTCTCTG 4682
 4622 GGGACCTGGAAGTGTGCTGCGACCCCAACAGCTTACTGATCAGCTGGAGTGTCTCTG 4681

4683 CTGTCAAGTGTAGATATTACAGGATCACTTACGGAGAAACAGGAGGAATAGCCCTCTCC 4742
 4682 CTGTCAAGTGTAGATATTACAGGATCACTTACGGAGAAACAGGAGGAATAGCCCTCTCC 4741
 4743 AGGAGTTTCACTGTGCTGGGAGCAAGTCTACAGTACCATCAGCGGCTTTAAACCTGGAG 4802
 4742 AGGAGTTTCACTGTGCTGGGAGCAAGTCTACAGTACCATCAGCGGCTTTAAACCTGGAG 4801
 4803 TTGATTTATACCATCACTGTGTATGCTGTCACTGGCGTGGAGACAGCCCGCAGAGCA 4862
 4802 TTGATTTATACCATCACTGTGTATGCTGTCACTGGCGTGGAGACAGCCCGCAGAGCA 4861
 4863 AGCCAAATTTCCATTAATTTACCCGAACAGAAATTTGACAAACCATCCAGATGCAAGTGA 4922
 4862 AGCCAAATTTCCATTAATTTACCCGAACAGAAATTTGACAAACCATCCAGATGCAAGTGA 4921
 4923 ATGTTTCAGGACAAACAGCATTTAGTGTCAAGTGGCTGCTTCAAGTTCCTGTTACTGGTT 4982
 4922 ATGTTTCAGGACAAACAGCATTTAGTGTCAAGTGGCTGCTTCAAGTTCCTGTTACTGGTT 4981
 4983 ACAGAGTTAAACCACTCCCAAAATGGACACAGGACCAACAAACAACTAAACCTGCAGTGC 5042
 4982 ACAGAGTTAAACCACTCCCAAAATGGACACAGGACCAACAAACAACTAAACCTGCAGTGC 5041
 5043 CAGATCAAAACAGAAATGACTATTGAAGCTTTGAGCCCAAGTGGAGTATGTTAGTGTG 5102
 5042 CAGATCAAAACAGAAATGACTATTGAAGCTTTGAGCCCAAGTGGAGTATGTTAGTGTG 5101
 5103 TCTATGCTCAGAAATCCAAAGCGAGAGAGTCAAGCTCTGTTTTCAGACTGCAGTAAACCA 5162
 5102 TCTATGCTCAGAAATCCAAAGCGAGAGAGTCAAGCTCTGTTTTCAGACTGCAGTAAACCA 5161
 5163 TTGATTCGCCCTAAAGGACTGCGATTTCATGTATGAGTGTGATTCATCAAAATTTGCTT 5222
 5162 TTGATTCGCCCTAAAGGACTGCGATTTCATGTATGAGTGTGATTCATCAAAATTTGCTT 5221
 5223 GGGAAAGCCCAAGGGCAAGTTTCCAGGTACAGGTCAGCTTCTCGAGCCCTGAGGATG 5282
 5222 GGGAAAGCCCAAGGGCAAGTTTCCAGGTACAGGTCAGCTTCTCGAGCCCTGAGGATG 5281
 5283 GAATCCATGAGCTATTTCCTGTCACCTGATGTTGAGAAAGACACTGCAGAGTGCAGAGGCC 5342
 5282 GAATCCATGAGCTATTTCCTGTCACCTGATGTTGAGAAAGACACTGCAGAGTGCAGAGGCC 5341
 5343 TCAGACCGGTTTCTGAGTACACAGTCACTGTGTTGCTTGCATGATATGAGAGGCC 5402
 5342 TCAGACCGGTTTCTGAGTACACAGTCACTGTGTTGCTTGCATGATATGAGAGGCC 5401
 5403 AGCCCTGATTTGGAACCCAGTCCACAGCTATTTCCTGACCAACTGACCTGAAGTTCACCTC 5462
 5402 AGCCCTGATTTGGAACCCAGTCCACAGCTATTTCCTGACCAACTGACCTGAAGTTCACCTC 5461
 5463 AGGTCAACCCCAAGGCTGAGCGCCCAAGTGGACACCAACCCCAATGTTTCACTGCTGAT 5522
 5462 AGGTCAACCCCAAGGCTGAGCGCCCAAGTGGACACCAACCCCAATGTTTCACTGCTGAT 5521
 5523 ATCGAGTGGGTTGACCCCAAGGAGAGACCGGACCAATGAAAGAAATCAACCTTTGCTC 5582
 5522 ATCGAGTGGGTTGACCCCAAGGAGAGACCGGACCAATGAAAGAAATCAACCTTTGCTC 5581
 5583 CTGACAGCTCATCCGCTGTTGATCAGGACTTATGGTGGCCACCAAAATATGAAAGTGTG 5642
 5582 CTGACAGCTCATCCGCTGTTGATCAGGACTTATGGTGGCCACCAAAATATGAAAGTGTG 5641
 5643 TCTATGCTCTTAAGGACACTTTTGAACAGCAGACAGTCTCAGGAGTGTGTCACCTCTGG 5702
 5642 TCTATGCTCTTAAGGACACTTTTGAACAGCAGACAGTCTCAGGAGTGTGTCACCTCTGG 5701
 5703 AGAATGTCTAGCCCAACCAAGAGAGGCTCGTGTGACAGATGCTTACTGAGACCAACCATCA 5762
 5702 AGAATGTCTAGCCCAACCAAGAGAGGCTCGTGTGACAGATGCTTACTGAGACCAACCATCA 5761

5763 TTAGCTGGAGAACCAAGACTGAGACGATCACTGGGCTCCAACTTGATGCCGTTCCAGCCA 5822
5762 TTAGCTGGAGAACCAAGACTGAGACGATCACTGGGCTCCAACTTGATGCCGTTCCAGCCA 5821
5823 ATGGCCAGAGCTCCAAATCCAGAGAACCAATCAAGCCAGATGTGAGAGCTACACCATCACTG 5882
5822 ATGGCCAGAGCTCCAAATCCAGAGAACCAATCAAGCCAGATGTGAGAGCTACACCATCACTG 5881
5883 GCTTACAAACAGGACCTGACTACAGATCTACTGCTGATGACCTTGATGACATGCTCGGA 5942
5882 GTTACAAACAGGACCTGACTACAGATCTACTGCTGATGACCTTGATGACATGCTCGGA 5941
5943 GCTCCCTGTGGTCACTGACGCTCCCACTGCCATTCATGACCATCCAACTCGCTTCC 6002
5942 GCTCCCTGTGGTCACTGACGCTCCCACTGCCATTCATGACCATCCAACTCGCTTCC 6001
6003 TGCCACACCAACCCAAATTCCTGCTGGTATCATGGCAGCGGCACTGTCAGGATTAACG 6062
6002 TGCCACACCAACCCAAATTCCTGCTGGTATCATGGCAGCGGCACTGTCAGGATTAACG 6061
6063 GCTACATCATCAAGTATGAGAGCTTGGTCTCTCCAGAGAGTGGTCCCTCGGCCCC 6122
6062 GCTACATCATCAAGTATGAGAGCTTGGTCTCTCCAGAGAGTGGTCCCTCGGCCCC 6121
6123 GCCTGGTGTCAAGAGGCTACTATTAATGCTGGCTGGAACCGGGAACCAATATCAATTT 6182
6122 GCCTGGTGTCAAGAGGCTACTATTAATGCTGGCTGGAACCGGGAACCAATATCAATTT 6181
6183 ATGCTATGCTCCCTGGAAGATTAATCAGAGAGGAGCCCTGATGGAAGGAAAGAGACAG 6242
6182 ATGCTATGCTCCCTGGAAGATTAATCAGAGAGGAGCCCTGATGGAAGGAAAGAGACAG 6241
6243 ACGAGTTCCTCCAACTGGTAAACCTTCCACACCCCAATCTTATGGAACCAAGATCTGG 6302
6242 ACGAGTTCCTCCAACTGGTAAACCTTCCACACCCCAATCTTATGGAACCAAGATCTGG 6301
6303 ATGCTTCCTCCAGTTCAGAGCCCTTCTGTCACCCACCCCTGGGTATGACCTGGAA 6362
6302 ATGCTTCCTCCAGTTCAGAGCCCTTCTGTCACCCACCCCTGGGTATGACCTGGAA 6361
6363 ATGCTATTCAGCTTCCTGGCACTTCTGGTACAGAACCCAGTGTGGGCAACAAATGATCT 6422
6362 ATGCTATTCAGCTTCCTGGCACTTCTGGTACAGAACCCAGTGTGGGCAACAAATGATCT 6421
6423 TTGAGGAACATGTTTATGAGGACCAACCGCCACACAGCCACCCCAATAGGCATA 6482
6422 TTGAGGAACATGTTTATGAGGACCAACCGCCACACAGCCACCCCAATAGGCATA 6481
6483 GGCCAAAGACCATACCCGCGGAATGTAGGACAGAGCTCTCTCTCAGACAAACCATCTCAT 6542
6482 GGCCAAAGACCATACCCGCGGAATGTAGGACAGAGCTCTCTCTCAGACAAACCATCTCAT 6541
6543 GGCCGCAATTCAGGACACTTCTGAGTACATCATTTTCATGTCATCTGTTGGCACTGATG 6602
6542 GGCCGCAATTCAGGACACTTCTGAGTACATCATTTTCATGTCATCTGTTGGCACTGATG 6601
6603 AAGAACCTTTACAGTTTCCAGGCTTCTACCACTTCTACCACTGCTGACAGGCTCA 6662
6602 AAGAACCTTTACAGTTTCCAGGCTTCTACCACTTCTACCACTGCTGACAGGCTCA 6661
6663 CAGAGTGCCACTTACCAATCATAGTGGAGCACTGGAAGACCAAGGAGGATAGG 6722
6662 CAGAGTGCCACTTACCAATCATAGTGGAGCACTGGAAGACCAAGGAGGATAGG 6721
6723 TTGGGAAGAGGTTTATCCGCTGGCACTCTGTCACCAAGAGCTTGAACCACTTACCG 6782
6722 TTGGGAAGAGGTTTATCCGCTGGCACTCTGTCACCAAGAGCTTGAACCACTTACCG 6781
6783 ATGACTCGTGTGTTGACCCCTACAGAGTCCCAATATGCGGTTGGAGATGAGTGGAAAC 6842
6782 ATGACTCGTGTGTTGACCCCTACAGAGTCCCAATATGCGGTTGGAGATGAGTGGAAAC 6841
6843 GAATGCTGTAATCAGGCTTTTAAACTGTTGTCAGGCTTCTAGGCTTTGGAAAGTGGTCAIT 6902

Db 6842 GAATGCTGTAATCAGGCTTTAAACTGTTGTCAGGCTCTAGGCTTTGGAAAGTGGTCAIT 6901
Qy 6903 TCAGATGTGATTCATCTAGATGTGTCATGACAAATGGTGTGAATCTAAGATTTGGAGAGA 6962
Db 6902 TCAGATGTGATTCATCTAGATGTGTCATGACAAATGGTGTGAATCTAAGATTTGGAGAGA 6961
Qy 6963 AGTGGGACCGTTCAGGAGAAATGGCCAGATGATGAGCTGCACATGCTTTGGAAACGGAA 7022
Db 6962 AGTGGGACCGTTCAGGAGAAATGGCCAGATGATGAGCTGCACATGCTTTGGAAACGGAA 7021
Qy 7023 AAGGAGAAATTCAGATGTGACCTCATGAGGCAACGTGTTACGATGATGGAAGACATACC 7082
Db 7022 AAGGAGAAATTCAGATGTGACCTCATGAGGCAACGTGTTACGATGATGGAAGACATACC 7081
Qy 7083 ACCTAGGAGAAACAGTGGCAGAAATATCTCGGTGCCATTTGCTCTCTGACATGCTTTG 7142
Db 7082 ACCTAGGAGAAACAGTGGCAGAAATATCTCGGTGCCATTTGCTCTCTGACATGCTTTG 7141
Qy 7143 GAGCCAGCGGGCTGGCGCTGTGCAAACTGCGCAGAACCTGGGGGTGAACCCAGTCCCG 7202
Db 7142 GAGCCAGCGGGCTGGCGCTGTGCAAACTGCGCAGAACCTGGGGGTGAACCCAGTCCCG 7201
Qy 7203 AAGCAGCTACTGSCCAGTCTCAACCCAGTATTCAGAGATACCATACAGAGAACAAACA 7262
Db 7202 AAGCAGCTACTGSCCAGTCTCAACCCAGTATTCAGAGATACCATACAGAGAACAAACA 7261
Qy 7263 CTAATGTTAAATGGCCAAATGAGTGTCTCATGCTTTAGATGTACAGGCTGACAGAGAG 7322
Db 7262 CTAATGTTAAATGGCCAAATGAGTGTCTCATGCTTTAGATGTACAGGCTGACAGAGAG 7321
Qy 7323 ATTCGAGAGATTAATCATCTTCCATCCAGAGAACCAAGCATGCTCTCTCCCAAGAT 7382
Db 7322 ATTCGAGAGATTAATCATCTTCCATCCAGAGAACCAAGCATGCTCTCTCCCAAGAT 7381
Qy 7383 CCATCTAAACTGGAGTGATGTAGCAGACCCAGCTTATAGTCTTTCTTTCTTTTAAAGC 7442
Db 7382 CCATCTAAACTGGAGTGATGTAGCAGACCCAGCTTATAGTCTTTCTTTCTTTTAAAGC 7441
Qy 7443 CTTTGTCTGGAGAGTCTCCAGCTTCAGCTCAACTCAGAGCTTCCCAAGCATCAC 7502
Db 7442 CTTTGTCTGGAGAGTCTCCAGCTTCAGCTCAACTCAGAGCTTCCCAAGCATCAC 7501
Qy 7503 CCTGGGAGTTTCTGAGGGTCTTCTATAAATGAGGCTGCACATTCGCTGTTCTGCTTC 7562
Db 7502 CCTGGGAGTTTCTGAGGGTCTTCTATAAATGAGGCTGCACATTCGCTGTTCTGCTTC 7561
Qy 7563 GAATATTCAATACCGCTCAGTATTTAATGAGTGAATTTAAGATTTGGTTGGGATC 7622
Db 7562 GAATATTCAATACCGCTCAGTATTTAATGAGTGAATTTAAGATTTGGTTGGGATC 7621
Qy 7623 AATAGGAAGCATATGAGGCAACCAAGATGCAAAATGTTTGAATGATATGACCAAAAT 7682
Db 7622 AATAGGAAGCATATGAGGCAACCAAGATGCAAAATGTTTGAATGATATGACCAAAAT 7681
Qy 7683 TTTAAGTAGGAAGTCAACCAACATTTCTGCTTTCTAATGATGCTGCGCCGCAATAC 7742
Db 7682 TTTAAGTAGGAAGTCAACCAACATTTCTGCTTTCTAATGATGCTGCGCCGCAATAC 7741
Qy 7743 TGTAGGAACAGCATGATCTTGTACTGTGATATTTTAAATATCCACAGTACT 7795
Db 7742 TGTAGGAACAGCATGATCTTGTACTGTGATATTTTAAATATCCACAGTACT 7794

RESULT 5
US-10-098-841-7
; Sequence 7, Application US/10098841
; Publication No. US20020197679A1
; GENERAL INFORMATION:
; APPLICANT: fang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Xu, Chongjun

APPLICANT: Zhou, Ping
 APPLICANT: Ma, Yunding
 APPLICANT: Wang, Jian-Rui
 APPLICANT: Zhao, Qing A.
 APPLICANT: Ren, Feiyan
 APPLICANT: Chen, Rui-hong
 APPLICANT: Wang, Dunrui
 APPLICANT: Wang, Zhiwei
 APPLICANT: Wehrman, Tom
 APPLICANT: Zhang, Jie
 APPLICANT: Qian, Xiaohong B.
 APPLICANT: Drmanac, Radjoje T.
 TITLE OF INVENTION: No. US20020197679A1el Nucleic Acids and
 TITLE OF INVENTION: Polypeptides
 FILE REFERENCE: 784C1P2
 CURRENT APPLICATION NUMBER: US/10/098,841
 CURRENT FILING DATE: 2002-03-13
 PRIOR APPLICATION NUMBER: 09/598,042
 PRIOR FILING DATE: 2000-06-20
 PRIOR APPLICATION NUMBER: 09/552,317
 PRIOR FILING DATE: 2000-04-25
 PRIOR APPLICATION NUMBER: 09/488,725
 PRIOR FILING DATE: 2000-01-21
 NUMBER OF SEQ ID NOS: 331
 SOFTWARE: pt_FL_genes Version 1.0
 SEQ ID NO 7
 LENGTH: 8230
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (363)..(7523)
 IS-10-098-841-7

Query Match 98.4%; Score 7671.8; DB 13; Length 8230;
 Best Local Similarity 98.7%; Pred. No. 0;
 Matches 7779; Conservative 0; Mismatches 7; Indels 93; Gaps 1;

2y	10	CCGCGCTGTGTCACAGGGGAGAGAGGAAACCCAGGCGGAGCGGGAAGAGGGAGC	69
2b	104	CCGCGCTGTGTCACAGGGGAGAGAGGAAACCCAGGCGGAGCGGGAAGAGGGAGC	163
2y	70	TGGAGCACAACTTCTGTGCTCTGCATCCCTTCTGCTCCCTCCAGCCGCTCCCTCC	129
2b	164	TGCAGCACAACTTCTGTGCTCTGCATCCCTTCTGCTCCCTCCAGCCGCTCCCTCC	223
2y	130	CACCTCTGGCCCCACCTTCTGGAGGCGACAAACCCCGGGAGGCAATTAGAAGGATTT	189
2b	224	CACCTCTGGCCCCACCTTCTGGAGGCGACAAACCCCGGGAGGCAATTAGAAGGATTT	283
2y	190	TTCCGCGAGGTGGAGAGGAGCAAACTTGGTGGCACTTGCCTCCCGTGGGGGCTC	249
2b	284	TTCCGCGAGGTGGAGAGGAGCAAACTTGGTGGCACTTGCCTCCCGTGGGGGCTC	343
Qy	250	TCTCCGCCACCGTCTCAACATGCTTAGGGGTCGGGGGCGCGGCTGCTGCTGGCGCT	309
Db	344	TCTCCGCCACCGTCTCAACATGCTTAGGGGTCGGGGGCGCGGCTGCTGCTGGCGCT	403
Qy	310	CCTGTGCTGGGACAGCGGTGCGCTTCAACAGGCGGCGGCGGCTGCTGCTGGCGCT	369
Db	404	CCTGTGCTGGGACAGCGGTGCGCTTCAACAGGCGGCGGCGGCTGCTGCTGGCGCT	463
Qy	370	GCAATGCTTCAAGCCCACTTCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	429
Db	464	GCAATGCTTCAAGCCCACTTCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	523
Qy	430	CAATGGAACCACTATCAGATAAATCAACAGTGGAGCGGACCTTACCTAGGCAATGCGTT	489
Db	524	CAATGGAACCACTATCAGATAAATCAACAGTGGAGCGGACCTTACCTAGGCAATGCGTT	583
Qy	490	GGTTTGTACTTGTATGAGGAGCGGAGGTTTAACTGCGAAAGTAAACCTGAAGCTGA	549
Db	584	GGTTTGTACTTGTATGAGGAGCGGAGGTTTAACTGCGAGAGTAAACCTGAAGCTGA	643

Qy	550	AGAGACTTGTCTTTCACAAAGTACACTGGGACACTTACCGAGTGGGTGACACTTATGAGCG	609
Db	644	AGAGACTTGTCTTTCACAAAGTACACTGGGACACTTACCGAGTGGGTGACACTTATGAGCG	703
Qy	610	TCCTAAAGACTTCCATGATCTGGGACTGTACCTGCATCGGGCTGGGCGAGGAGATAAG	669
Db	704	TCCTAAAGACTTCCATGATCTGGGACTGTACCTGCATCGGGCTGGGCGAGGAGATAAG	763
Qy	670	CTGTACCATGCAAAACCGCTGCCATGAGGGGCTGAGTCCCTCAAGATTGGTGACACCTG	729
Db	764	CTGTACCATGCAAAACCGCTGCCATGAGGGGCTGAGTCCCTCAAGATTGGTGACACCTG	823
Qy	730	GAGGAGACCAATGAGACTGCTGTGTACATGTTAGAGTGTGTGTCTTGGTAAATGAAA	789
Db	824	GAGGAGACCAATGAGACTGCTGTGTACATGTTAGAGTGTGTGTCTTGGTAAATGAAA	883
Qy	790	AGGAGATGGACCTGCAAGCCCATAGCTGAGAGTGTGTTTATGATCATGCTCTGGGACTTC	849
Db	884	AGGAGATGGACCTGCAAGCCCATAGCTGAGAGTGTGTTTATGATCATGCTCTGGGACTTC	943
Qy	850	CTATGTGTGCGAGAAAACGCTGGGAGAAAGCCCTACCAAGGCTGGATGATGTTAGATTGAC	909
Db	944	CTATGTGTGCGAGAAAACGCTGGGAGAAAGCCCTACCAAGGCTGGATGATGTTAGATTGAC	1003
Qy	910	TGCGTGGGAGAGGCGAGCGGACCATCTTGCATCTTAGAAATAGATGCAACGATCA	969
Db	1004	TGCGTGGGAGAGGCGAGCGGACCATCTTGCATCTTAGAAATAGATGCAACGATCA	1063
Qy	970	GGACACAGGACATCCCTATAGAAATGGAGACACCTGGAGCAAGAGGATAATCCAGGAAA	1029
Db	1064	GGACACAGGACATCCCTATAGAAATGGAGACACCTGGAGCAAGAGGATAATCCAGGAAA	1123
Qy	1030	CCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGAGAGTGGAAAGTGTGAGAGGCAC	1089
Db	1124	CCTGCTCCAGTGCATCTGCAACAGCAACGCGCGAGGAGAGTGGAAAGTGTGAGAGGCAC	1183
Qy	1090	CTCTGTGAGACCCATCATGAGCGGATCTGGCCCTTCCACCGATGTTGTCAGCTGTTTA	1149
Db	1184	CTCTGTGAGACCCATCATGAGCGGATCTGGCCCTTCCACCGATGTTGTCAGCTGTTTA	1243
Qy	1150	CCAAACCGAGCCTCACACCCCGAGCCTCTCTTATGGCCACTGTGTACAGACAGTGTGT	1209
Db	1244	CCAAACCGAGCCTCACACCCCGAGCCTCTCTTATGGCCACTGTGTACAGACAGTGTGT	1303
Qy	1210	GGTCTACTCTGTGGGATGAGTGGCTGAAGACACAGGAAATAGCAAAATGCTTTGCAC	1269
Db	1304	GGTCTACTCTGTGGGATGAGTGGCTGAAGACACAGGAAATAGCAAAATGCTTTGCAC	1363
Qy	1270	GTGCGTGGGCAACGAGTCACTGTCACAGACAGCTGTAAACCCAGACTTACGGTGGCAA	1329
Db	1364	GTGCGTGGGCAACGAGTCACTGTCACAGACAGCTGTAAACCCAGACTTACGGTGGCAA	1423
Qy	1330	CTCAATGGAGAGCAATGTGTCTTACCATTCACCTAGATGGCAGGAGTCTACTCTCTG	1389
Db	1424	CTCAATGGAGAGCAATGTGTCTTACCATTCACCTAGATGGCAGGAGTCTACTCTCTG	1483
Qy	1390	CACCAAGAGGCGGACAGGAGCGGACATCTTGGTGCAGCACAACCTTCAATATGAGCA	1449
Db	1484	CACCAAGAGGCGGACAGGAGCGGACATCTTGGTGCAGCACAACCTTCAATATGAGCA	1543
Qy	1450	GGACAGGAAATACCTTCTGTCACAGACCACTGTTTGGTTCAGACTCGAGGAGGAAA	1509
Db	1544	GGACAGGAAATACCTTCTGTCACAGACCACTGTTTGGTTCAGACTCGAGGAGGAAA	1603
Qy	1510	TTCCCATGTGCTTGTGCGACTTCCCTTCTTATCAACCAACCAAAATACACTGATTG	1569
Db	1604	TTCCCATGTGCTTGTGCGACTTCCCTTCTTATCAACCAACCAAAATACACTGATTG	1663
Qy	1570	CACTTCTGAGGCGAGAGAGCAACATGAAGTGTGTGGGACCAACAGAACTATGATGC	1629
Db	1664	CACTTCTGAGGCGAGAGAGCAACATGAAGTGTGTGGGACCAACAGAACTATGATGC	1723

1630 CGACAGAGTTGGTCTGCCCCCATGGCTGCCACGAGGAAATCTGCACAAACCAATGA 1689
1724 CCACAGAGAGTTGGTCTGCCCCCATGGCTGCCACGAGGAAATCTGCACAAACCAATGA 1783
1690 AGGGGTCAATGTACCGCATTTGGAGATCAGTGGGATAGCAGCATGATGAGTGCATGAT 1749
1784 AGGGGTCAATGTACCGCATTTGGAGATCAGTGGGATAGCAGCATGATGAGTGCATGAT 1843
1750 GAGTGCAGCTGTGTGGAAATGGTGGGAAATGGGATGATGATGATGATGATGATGATGAT 1809
1844 GAGTGCAGCTGTGTGGAAATGGTGGGAAATGGGATGATGATGATGATGATGATGATGAT 1903
1810 TCGAGATCAGTGCATTTGTTGATGATCAGTGCATTTGATGATGATGATGATGATGATGAT 1869
1904 TCGAGATCAGTGCATTTGTTGATGATCAGTGCATTTGATGATGATGATGATGATGATGAT 1963
1870 TCATGAGAGGGGACATGCTGAACTGTATCATGCTTGGTTCAGGTCGGGTCAGGTCGAA 1929
1964 TCATGAGAGGGGACATGCTGAACTGTATCATGCTTGGTTCAGGTCGGGTCAGGTCGAA 2023
1930 GTGTGATCCCGTCGACCAATGCCAGGATTCAGAGACTGGGACGTTTATCAAAATGGAGA 1989
2024 GTGTGATCCCGTCGACCAATGCCAGGATTCAGAGACTGGGACGTTTATCAAAATGGAGA 2083
1990 TTTCATGGGAGAGATGTCATGCTGGTTCAGATACAGTCTACTGCTATGGCCGTCGAT 2049
2084 TTTCATGGGAGAGATGTCATGCTGGTTCAGATACAGTCTACTGCTATGGCCGTCGAT 2143
2050 TGGGAGTGGGATTCGCAACCTTTACAGACCTATCCAACTCAAGTGTGCTGTGCAAGT 2109
2144 TGGGAGTGGGATTCGCAACCTTTACAGACCTATCCAACTCAAGTGTGCTGTGCAAGT 2203
2110 ATTATATCAGTACGAGTCCGAGTCAGGTCAGGTCAGTCTACTGCTATGGCCGTCGAT 2169
2204 ATTATATCAGTACGAGTCCGAGTCAGGTCAGGTCAGTCTACTGCTATGGCCGTCGAT 2263
2170 GCCATCTCAATTTCAAAGTACATTTCTCAGTGGGACCTAAAAATTTCTGTAGGCGGTG 2229
2264 GCCATCTCAATTTCAAAGTACATTTCTCAGTGGGACCTAAAAATTTCTGTAGGCGGTG 2323
2230 GAAGGAGTACCATACAGGCGGACCTTAACTCTTACATCAGTCAAGGCTTGAAGCTGG 2289
2324 GAAGGAGTACCATACAGGCGGACCTTAACTCTTACATCAGTCAAGGCTTGAAGCTGG 2383
2290 TGTGGTATACGAGGCGGACCTCATAGCATCCAGCATGAGGTCAGGTCAGGTCAGGTCAG 2349
2384 TGTGGTATACGAGGCGGACCTCATAGCATCCAGCATGAGGTCAGGTCAGGTCAGGTCAG 2443
2350 CTTTGTACTTCAACCAACCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2409
2444 CTTTGTACTTCAACCAACCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2503
2410 GAGACTCCCTTTTCTCTCTGTCGCTTCTGTCGCTTCTGTCGCTTCTGTCGCTTCTGTC 2469
2504 GAGACTCCCTTTTCTCTCTGTCGCTTCTGTCGCTTCTGTCGCTTCTGTCGCTTCTGTC 2563
2470 TAGCTTTGTGGTCTCCTGGTCTCAGCTTCCGACACCGTGTGGGATTCGGGTGGAATA 2529
2564 TAGCTTTGTGGTCTCCTGGTCTCAGCTTCCGACACCGTGTGGGATTCGGGTGGAATA 2623
2530 TGAGCTGAGTGGAGGAGATCAGCAGCAGTACCTGGATCTTCCAGCAGCAGCAGCAGCAG 2589
2624 TGAGCTGAGTGGAGGAGATCAGCAGCAGTACCTGGATCTTCCAGCAGCAGCAGCAGCAG 2683
2590 TGTGAACATCCCTGACCTGCTTCTGTCGCTTCTGTCGCTTCTGTCGCTTCTGTCGCTTCT 2649
2684 TGTGAACATCCCTGACCTGCTTCTGTCGCTTCTGTCGCTTCTGTCGCTTCTGTCGCTTCT 2743
2650 TGAGGATGGGAGCAGATTTGATCTCTGTCATCTTCAACACACAGGCTGATGCCCT 2709
2744 TGAGGATGGGAGCAGATTTGATCTCTGTCATCTTCAACACACAGGCTGATGCCCT 2803
2710 TCCTGACCCGACTGTGGACCAAGTTGATGACACTCAATTTGTTTCTGCTGGAGCAGACC 2769

2804 TCCTGACCCGACTGTGGACCAAGTTGATGACACCTCAATTTGTTGCTGGAGCAGACC 2863
2770 CCAGGCTCCCATCACAGGCTACAGATAGTCTATTTCGCCATCAGTAGAGGTAGCAGCAC 2829
2864 CCAGGCTCCCATCACAGGCTACAGATAGTCTATTTCGCCATCAGTAGAGGTAGCAGCAC 2923
2830 AGAACTCAACCTTCTGAAAACCTGCAAACTCCGTCACCTCAGTGAATTTGCAACCTGGTGT 2889
2924 AGAACTCAACCTTCTGAAAACCTGCAAACTCCGTCACCTCAGTGAATTTGCAACCTGGTGT 2983
2890 TCAGTATACATCAGTCTATTTCGCTGTGGAAGAAATCAAGAAAGTACACCTGTTGTCTAT 2949
2984 TCAGTATACATCAGTCTATTTCGCTGTGGAAGAAATCAAGAAAGTACACCTGTTGTCTAT 3043
2950 TCAACAGAAACCACTGGCACCCCGCTCAGATACAGTGCCTCTTCCAGGACCTGTCA 3009
3044 TCAACAGAAACCACTGGCACCCCGCTCAGATACAGTGCCTCTTCCAGGACCTGTCA 3103
3010 GTTGTGGAGTGCAGAGCGTGAAGTCAACATCATGTGACACCGCTTGAAGTGCAGT 3069
3104 GTTGTGGAGTGCAGAGCGTGAAGTCAACATCATGTGACACCGCTTGAAGTGCAGT 3163
3070 GACCGGCTACCGTGTGATGATCCCGCTCAACCTGCGGTCGCGAGCACGGGACAGGCT 3129
3164 GACCGGCTACCGTGTGATGATCCCGCTCAACCTGCGGTCGCGAGCACGGGACAGGCT 3223
3130 GCCCATCAGCAGGAAACCTTTTCAGAGAGTCAACGCGCTGTCCCTGGGTCACCTATTAT 3189
3224 GCCCATCAGCAGGAAACCTTTTCAGAGAGTCAACGCGCTGTCCCTGGGTCACCTATTAT 3283
3190 CTTCAAAGTCTTTGCACTGAGCCATGGGAGGAGAGCAAGCTCTGACTGCTCAACAGAC 3249
3284 CTTCAAAGTCTTTGCACTGAGCCATGGGAGGAGAGCAAGCTCTGACTGCTCAACAGAC 3343
3250 AACCAACTGATGCTCCCACTAACCTCCAGTTTGTCAATGAACCTGATTTCTACTGCTCT 3309
3344 AACCAACTGATGCTCCCACTAACCTCCAGTTTGTCAATGAACCTGATTTCTACTGCTCT 3403
3310 GGTGAGATGAGTCCCACTCGGCGCCAGATAACAGGATACCGACTGACCGTGGGCGCTTAC 3369
3404 GGTGAGATGAGTCCCACTCGGCGCCAGATAACAGGATACCGACTGACCGTGGGCGCTTAC 3463
3370 CCGAAGAGGACAGCCCAAGGAGTACAAATGTGGTTCCTCTGTCTCCAGTACCCTGAG 3429
3464 CCGAAGAGGACAGCCCAAGGAGTACAAATGTGGTTCCTCTGTCTCCAGTACCCTGAG 3523
3430 GAATCTGAGCTGCTGATGATACCTGATCCCTCGTGGCCATTAAGGGCAACCAAGA 3489
3524 GAATCTGAGCTGCTGATGATACCTGATCCCTCGTGGCCATTAAGGGCAACCAAGA 3583
3490 GAGCCCAAGGACCTGAGTCTTTTACCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3549
3584 GAGCCCAAGGACCTGAGTCTTTTACCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3643
3550 CAACCCGAGGAGTGCAGACCCCTTGTGATCAGATGAGGAGGAGGAGGAGGAGGAGGAGGAG 3609
3644 CAACCCGAGGAGTGCAGACCCCTTGTGATCAGATGAGGAGGAGGAGGAGGAGGAGGAGGAG 3703
3610 TTTTAAGCTGGGTGTACGACCAAGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3669
3704 TTTTAAGCTGGGTGTACGACCAAGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3763
3670 CTCAGGAGGATCTGTTGTCGGCTTGTGATCAGATGAGGAGGAGGAGGAGGAGGAGGAGGAG 3729
3764 CTCAGGAGGATCTGTTGTCGGCTTGTGATCAGATGAGGAGGAGGAGGAGGAGGAGGAGGAG 3823
3730 AGTCTGAGAGATGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3789
3824 AGTCTGAGAGATGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3883
3790 GTCTCCACCAACAACTTGTGATCTGGAGGCAACCTTGAACCTGAGCTGGAGTGTCTCAGTCTC 3849

Db 3884 GTCTCCACCAAACTTGGCATCTGGAGGCAAAACCTGCACACCTGGAGTGTCTACAGTCTC 3943
 Qy 3850 CTGGAGAGGAGCAACCCAGACATTAATCTGGTTATAGAAATACCAACACCCCTACAAA 3909
 Db 3944 CTGGAGAGGAGCAACCCAGACATTAATCTGGTTATAGAAATACCAACACCCCTACAAA 4003
 Qy 3910 CGGCACAGGAGAAATCTTTTGAAGAAGTGGTCCATCTGCTGATCAGAGCTCTCGCACTTT 3969
 Db 4004 CGGCACAGGAGAAATCTTTTGAAGAAGTGGTCCATCTGCTGATCAGAGCTCTCGCACTTT 4063
 Qy 3970 TGATAAATCTGAGTCCGGCTCTGGAGTACAAATGTCAGTGTGTATACCTGTCAAGATGACAA 4029
 Db 4064 TGATAAATCTGAGTCCGGCTCTGGAGTACAAATGTCAGTGTGTATACCTGTCAAGATGACAA 4123
 Qy 4030 GGAAGTGTCCCTATCTCTGATACCATCATCCAGCTGTTCCTCTCCCTCCACTGACCTGGC 4089
 Db 4124 GGAAGTGTCCCTATCTCTGATACCATCATCCAGCTGTTCCTCTCCCTCCACTGACCTGGC 4183
 Qy 4090 ATTCAACCAATTTGGTCAGACACCATGCTGTGTCACTGGGCTCCACCCCATTCATTTGA 4149
 Db 4184 ATTCAACCAATTTGGTCAGACACCATGCTGTGTCACTGGGCTCCACCCCATTCATTTGA 4243
 Qy 4150 TTTAAACCAATTTCTGGTGGCTTACTCACTGTGAAATAAGGAGATGTTGACAGATT 4209
 Db 4244 TTTAAACCAATTTCTGGTGGCTTACTCACTGTGAAATAAGGAGATGTTGACAGATT 4303
 Qy 4210 GTCAATTTCTCTTTCAGACAAATGAGTGTGTTTAAACAAATCTCTGCTGGTACAGAAAT 4269
 Db 4304 GTCAATTTCTCTTTCAGACAAATGAGTGTGTTTAAACAAATCTCTGCTGGTACAGAAAT 4363
 Qy 4270 TGTAGTGTGCTCTCCAGTGTCTAGACACATGAGAGACACCTCTTAGAGAGACACA 4329
 Db 4364 TGTAGTGTGCTCTCCAGTGTCTAGACACATGAGAGACACCTCTTAGAGAGACACA 4423
 Qy 4330 GAAAAACAGTCTTGATTTCCCAACTGGCAATTTCTGATATTTACTGCCAATCTTTT 4389
 Db 4424 GAAAAACAGTCTTGATTTCCCAACTGGCAATTTCTGATATTTACTGCCAATCTTTT 4483
 Qy 4390 TACTGTGACGTGGATGCTCTCGAGCCACCATCACTGGCTACAGGATCGCCATCATCC 4449
 Db 4484 TACTGTGACGTGGATGCTCTCGAGCCACCATCACTGGCTACAGGATCGCCATCATCC 4543
 Qy 4450 CGAGCACTTCAGTGGAGACCTCTCAGAGAGATCGGGTGCCCACTCTCGGAATTCATCAC 4509
 Db 4544 CGAGCACTTCAGTGGAGACCTCTCAGAGAGATCGGGTGCCCACTCTCGGAATTCATCAC 4603
 Qy 4510 CCTCACCAACTCTCACTCAGGACACAGATATGCTGACGATCGTTCTTTATGGCAG 4569
 Db 4604 CCTCACCAACTCTCACTCAGGACACAGATATGCTGACGATCGTTCTTTATGGCAG 4663
 Qy 4570 AGAGAAAGTCCCTTATTTGATTTGGCCAAATCAACAGTTTCTGATGTTCCGAGGGACCT 4629
 Db 4664 AGAGAAAGTCCCTTATTTGATTTGGCCAAATCAACAGTTTCTGATGTTCCGAGGGACCT 4723
 Qy 4630 GGAAGTGTGCTGGACCCCAACAGCCTTCTGATGCTGGATGCTCTGCTGTTCAC 4689
 Db 4724 GGAAGTGTGCTGGACCCCAACAGCCTTCTGATGCTGGATGCTCTGCTGTTCAC 4783
 Qy 4690 AGTGAGATTTACAGATCACTTACGGAGAAACAGAGGAAATAGCCTGTCCAGGAGTT 4749
 Db 4784 AGTGAGATTTACAGATCACTTACGGAGAAACAGAGGAAATAGCCTGTCCAGGAGTT 4843
 Qy 4750 CACTGTGCTGGAGCAAGTCTTACAGTACCATCAGCGCCCTTAAACCTGGAGTTGATTA 4809
 Db 4844 CACTGTGCTGGAGCAAGTCTTACAGTACCATCAGCGCCCTTAAACCTGGAGTTGATTA 4903
 Qy 4810 TACCATCACTGTGTATGCTGCTCACTGGCCGTGGAGACAGCCCGCAGCAGCAAGCCAAAT 4869
 Db 4904 TACCATCACTGTGTATGCTGCTCACTGGCCGTGGAGACAGCCCGCAGCAGCAAGCCAAAT 4963
 Qy 4870 TTCCATTAATACCGAAACAGAAATTTGACAAACCATCCAGATGCAAGTGACCGATGTTCA 4929
 Db 4964 TTCCATTAATACCGAAACAGAAATTTGACAAACCATCCAGATGCAAGTGACCGATGTTCA 5023

Qy 4930 GGACAAACAGCATTTAGTGTCAAGTGGCTGCTTCAAGTTCCCTCTGTTACTGGTTACAGAGT 4989
 Db 5024 GGACAAACAGCATTTAGTGTCAAGTGGCTGCTTCAAGTTCCCTCTGTTACTGGTTACAGAGT 5083
 Qy 4990 AACCCACCTCCCAAAATAGGACAGGACCAACAAATAAACTAAACTGACAGTCCAGATCA 5049
 Db 5084 AACCCACCTCCCAAAATAGGACAGGACCAACAAATAAACTAAACTGACAGTCCAGATCA 5143
 Qy 5050 AACAGAAATGACTATTGAAAGGCTTCAGCCCAACAGTGGAGTATGTGGTTAGTGTCTATGC 5109
 Db 5144 AACAGAAATGACTATTGAAAGGCTTCAGCCCAACAGTGGAGTATGTGGTTAGTGTCTATGC 5203
 Qy 5110 TCAGAAATCCAAACGAGAGAGTCACTCTCTGTTTCACTGAGTAAACCAATTTGATCG 5169
 Db 5204 TCAGAAATCCAAACGAGAGAGTCACTCTCTGTTTCACTGAGTAAACCAATTTGATCG 5263
 Qy 5170 CCTTAAAGGACTGGCAATTCACATGATGTGATTCGATTCATCAAAATTTGCTTGGGAAAG 5229
 Db 5264 CCTTAAAGGACTGGCAATTCACATGATGTGATTCGATTCATCAAAATTTGCTTGGGAAAG 5323
 Qy 5230 CCCACAGGGGCAAGTTTCCAGGTACAGGTTGACCTTACGAGCCCTGAGGATGGAATCCA 5289
 Db 5324 CCCACAGGGGCAAGTTTCCAGGTACAGGTTGACCTTACGAGCCCTGAGGATGGAATCCA 5383
 Qy 5290 TGAGCTATTTCCTGCACTGTGTTGAGAGACACTGAGAGCTGCAAGGCTCAGACC 5349
 Db 5384 TGAGCTATTTCCTGCACTGTGTTGAGAGACACTGAGAGCTGCAAGGCTCAGACC 5443
 Qy 5350 GGGTTCTGAGTACACAGTCACTGTTGCTTGCCTTGACGATGATATGGAGAGCCAGCCCT 5409
 Db 5444 GGGTTCTGAGTACACAGTCACTGTTGCTTGCCTTGACGATGATATGGAGAGCCAGCCCT 5503
 Qy 5410 GATTGAAACCCAGTCCAAGCTATTTCCTGACCACTGACCTGAAAGTTCACTCAGGTCA 5469
 Db 5504 GATTGAAACCCAGTCCAAGCTATTTCCTGACCACTGACCTGAAAGTTCACTCAGGTCA 5563
 Qy 5470 ACCCAACAGCTGAGCGCCAGTGACACCACTTATGTTGAGTGTGAGTGTGATGCTATGC 5529
 Db 5564 ACCCAACAGCTGAGCGCCAGTGACACCACTTATGTTGAGTGTGAGTGTGATGCTATGC 5623
 Qy 5530 GGGGTGACCCCAAGGAGAACCGGACCAATGAAAGAAATCAACCTTGTCTCTGACAG 5589
 Db 5624 GGGGTGACCCCAAGGAGAACCGGACCAATGAAAGAAATCAACCTTGTCTCTGACAG 5683
 Qy 5590 CTATCCGTTGTTGATCAGGACTTATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5649
 Db 5684 CTATCCGTTGTTGATCAGGACTTATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5743
 Qy 5650 TCTTAAAGGACACTTTGACAAAGCAGACCTCAGGAGTGTGTCACCACTCTGGAGAAATGT 5709
 Db 5744 TCTTAAAGGACACTTTGACAAAGCAGACCTCAGGAGTGTGTCACCACTCTGGAGAAATGT 5803
 Qy 5710 CAGCCCAACAGAGGGCTGCTGACAGATCTGAGACCACTCAGACCACTCAGCATTAGCTG 5769
 Db 5804 CAGCCCAACAGAGGGCTGCTGACAGATCTGAGACCACTCAGACCACTCAGCATTAGCTG 5863
 Qy 5770 GAGAACCAAGACTGAGACGATCACTGCTGCTCAGTGTGATGCTGCTGCTGCTGCTGCTGCT 5829
 Db 5864 GAGAACCAAGACTGAGACGATCACTGCTGCTCAGTGTGATGCTGCTGCTGCTGCTGCTGCT 5923
 Qy 5830 GACTCCAATCCAGAGAACCATCAGCCAGATCTCAGAGCTCAGACCACTCAGCATTAGCTG 5889
 Db 5924 GACTCCAATCCAGAGAACCATCAGCCAGATCTCAGAGCTCAGACCACTCAGCATTAGCTG 5983
 Qy 5890 ACCAGGACTGACTCAGAGATCTTACCTGTACACCTTGAATGCAATGTCTGGAGTCTCCC 5949
 Db 5984 ACCAGGACTGACTCAGAGATCTTACCTGTACACCTTGAATGCAATGTCTGGAGTCTCCC 6043
 Qy 5950 TGTGTTCTGACGCTCTCAGTGCATTTGATGCACTTCAACCTGCTGCTGCTGCTGCTGCT 6009
 Db 6044 TGTGTTCTGACGCTCTCAGTGCATTTGATGCACTTCAACCTGCTGCTGCTGCTGCTGCTGCT 6103

6010 CACACCAATTCCTTCTGGTATCATGGCAGCCGACGTCGAGATTACCGGCTACAT 6069
 6104 CACACCAATTCCTTCTGGTATCATGGCAGCCGACGTCGAGATTACCGGCTACAT 6163
 6070 CATCAAGTATGAGAAGCTGGGTCTCTCCAGAGAAGTGGTCCCTCGGCCCGCCCTGG 6129
 6164 CATCAAGTATGAGAAGCTGGGTCTCTCCAGAGAAGTGGTCCCTCGGCCCGCCCTGG 6223
 6130 TGTACAGAGGCTACTATTACTGGCTGGAACCCGGGAACCGGAATATACAAATTTATGTCAT 6189
 6224 TGTACAGAGGCTACTATTACTGGCTGGAACCCGGGAACCGGAATATACAAATTTATGTCAT 6283
 6190 TGCCCTGAAGAATAATACAGAGAGGAGCCCTGTATGGAAGGAAGGAAGAGACAGAGCT 6249
 6284 TGCCCTGAAGAATAATACAGAGAGGAGCCCTGTATGGAAGGAAGGAAGAGACAGAGCT 6343
 6250 TGCCCAAGCTGAACCCCTTCCACACCCCAATCTTCATGGACAGAGATCTTGGATGTTCC 6309
 6344 TGCCCAAGCTGAACCCCTTCCACACCCCAATCTTCATGGACAGAGATCTTGGATGTTCC 6403
 6310 TTCCACAGTTCAAAAAGACCCCTTTTGTACACCACTGGGTAATGACATGGAATGGTAT 6369
 6404 TTCCACAGTTCAAAAAGACCCCTTTTGTACACCACTGGGTAATGACATGGAATGGTAT 6463
 6370 TCAGCTTCTGGCACTCTCTGTCAGCAACCCAGTGTGGGCAACAAATGATCTTTGAGGA 6429
 6464 TCAGCTTCTGGCACTCTCTGTCAGCAACCCAGTGTGGGCAACAAATGATCTTTGAGGA 6523
 6430 ACATGTTTTAGCGGACACACACCCGCCCAACAGGCCACCCCAATAGGCATAGGCCAAG 6489
 6524 ACATGTTTTAGCGGACACACACCCGCCCAACAGGCCACCCCAATAGGCATAGGCCAAG 6583
 6490 ACCATACCGCCGAATGT----- 6507
 6584 ACCATACCGCCGAATGTAGGTAGGAAATCCAAATGTTGTACATCCCGAGGAAGATGT 6643
 6508 -----AGGCAAGA 6516
 6644 AGACTATACCTGTACCCACACGTCGGGGTCAATCCAAATGCTCTACAGGACAGA 6703
 6517 AGCTCTCTCAGACAAACATCTCATGGGCCCATTCACAGGACACTCTCAGTATCATAT 6576
 6704 AGCTCTCTCAGACAAACATCTCATGGGCCCATTCACAGGACACTCTCAGTATCATAT 6763
 6577 TTCATGTCATCTCTGGCACTCATGAAGAACCTTTACAGTTTCAGGGTTCCTGGAACCTTC 6636
 6764 TTCATGTCATCTCTGGCACTCATGAAGAACCTTTACAGTTTCAGGGTTCCTGGAACCTTC 6823
 6637 TACAGTGCCATCTCAGACAGGCTCACCAGAGTGCCACCTACACATCATATAGTGAGGC 6696
 6824 TACAGTGCCATCTCAGACAGGCTCACCAGAGTGCCACCTACACATCATATAGTGAGGC 6883
 6697 ACTGAAGACACAGAGGATAGGTTTGGGAAGAGGTTGTTACCGTGGGCAACTCTGT 6756
 6884 ACTGAAGACACAGAGGATAGGTTTGGGAAGAGGTTGTTACCGTGGGCAACTCTGT 6943
 6757 CAACGAAGGTTGAAACCAACTACAGATGACTCGTGTGAGCCCTACACAGTGTCCTCA 6816
 6944 CAACGAAGGTTGAAACCAACTACAGATGACTCGTGTGAGCCCTACACAGTGTCCTCA 7003
 6817 TTATGCGTTGGAGATGAGTGGGAACGAATGCTGTAATCAGGCTTTAAACTGTTGTCGA 6876
 7004 TTATGCGTTGGAGATGAGTGGGAACGAATGCTGTAATCAGGCTTTAAACTGTTGTCGA 7063
 6877 GTGCTTAGGCTTTGGAAAGTGGTCAATTCAGATGATGATTCATATGATGTTGTCGCAAA 6936
 7064 GTGCTTAGGCTTTGGAAAGTGGTCAATTCAGATGATGATTCATATGATGTTGTCGCAAA 7123
 6937 TGGTGTCAACTACAAAGTTGAGAGAGTGGGACCGTCAGGAGGAATAATGGCCAGATGAT 6996
 7124 TGGTGTCAACTACAAAGTTGAGAGAGTGGGACCGTCAGGAGGAATAATGGCCAGATGAT 7183
 6997 GAGCTGCACATGTTCTGGGAACGGAAAGAGGAATTCAAAGTGTGACCCCTCATGAGGCAAC 7056

7184 GAGCTGCACATGTCTTGGGAACGGAAGAGAAATTCAGTGTGACCTCTCATGAGGCAAC 7243
 7057 GTGTTACGATGATGGAAGACATACACCTAGGAGAACAGTGGCAGAAAGGAATATCTCGG 7116
 7244 GTGTTACGATGATGGAAGACATACACCTAGGAGAACAGTGGCAGAAAGGAATATCTCGG 7303
 7117 TGCATTTGCTCTGTCACATGCTTTGGAGGCGCAGCGGGCTGGCGCTGACAACTGGCG 7176
 7304 TGCATTTGCTCTGTCACATGCTTTGGAGGCGCAGCGGGCTGGCGCTGACAACTGGCG 7363
 7177 CAGACCTGGGGGTGAACCCAGTCCCGAAGGCACTACTGGCCAGTCTCAACACAGATATTC 7236
 7364 CAGACCTGGGGGTGAACCCAGTCCCGAAGGCACTACTGGCCAGTCTCAACACAGATATTC 7423
 7237 TCAGAGATACCATCAGAGAACAAACCTAATGTTAATTCGCCAATTTGAGTGTCTCATGCC 7296
 7424 TCAGAGATACCATCAGAGAACAAACCTAATGTTAATTCGCCAATTTGAGTGTCTCATGCC 7483
 7297 TTTAGATGTACAGGCTGACAGAGAAATTCCTGAGAGTAAATCATCTTTCCAAATCCAGAG 7356
 7484 TTTAGATGTACAGGCTGACAGAGAAATTCCTGAGAGTAAATCATCTTTCCAAATCCAGAG 7543
 7357 GAACAAGCATGTCTCTGTCGCAAGATCCATCTAATCTGAGTGTCTTAGCAGACCCAGC 7416
 7544 GAACAAGCATGTCTCTGTCGCAAGATCCATCTAATCTGAGTGTCTTAGCAGACCCAGC 7603
 7417 TTAGAGTCTCTCTCTTCTTAAAGCCCTTTGCTCTGAGGAGTCTCCAGCTTCAGCT 7476
 7604 TTAGAGTCTCTCTCTTCTTAAAGCCCTTTGCTCTGAGGAGTCTCCAGCTTCAGCT 7663
 7477 CAATCTACAGCTTCTCCAGCATCACCTGGGAGTTTCCTGAGGGTTTCTCTATAAATGA 7536
 7664 CAATCTACAGCTTCTCCAGCATCACCTGGGAGTTTCTCTGAGGGTTTCTCTATAAATGA 7723
 7537 GGGCTGCACATTCGCTGTTCTGCTTGAAGTATTCATACCGCTCAGTATTTTAAATGAA 7596
 7724 GGGCTGCACATTCGCTGTTCTGCTTGAAGTATTCATACCGCTCAGTATTTTAAATGAA 7783
 7597 GTGATTTCTAAGATTTGGTATGGGATCAATAGGAAGCATATGAGCCAAACCAAGATGCAA 7656
 7784 GTGATTTCTAAGATTTGGTATGGGATCAATAGGAAGCATATGAGCCAAACCAAGATGCAA 7843
 7657 ATGTTTTGAATGATGACCAAAATTTTAAAGTAGGAAGTCAACCAACACTCTCTGCTT 7716
 7844 ATGTTTTGAATGATGATGACCAAAATTTTAAAGTAGGAAGTCAACCAACACTCTCTGCTT 7903
 7717 TCATTAAGTGTCTGGCCCGCAATCTGTTAGGAACAGCATGATCTTGTACTGTGATAT 7776
 7904 TCATTAAGTGTCTGGCCCGCAATCTGTTAGGAACAGCATGATCTTGTACTGTGATAT 7963
 7777 TTTAAATATCCAGTACT 7795
 7964 TTTAAATATCCAGTACT 7982

RESULT 6
 US-10-098-841-5
 ; Sequence 5, Application US/10098841
 ; Publication No. US20020197679A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Liu, Chenghua
 ; APPLICANT: Asundi, Vinod
 ; APPLICANT: Xu, Chongjun
 ; APPLICANT: Zhou, Ping
 ; APPLICANT: Wang, Jian-Rui
 ; APPLICANT: Zhao, Qing A.
 ; APPLICANT: Ren, Feiyan
 ; APPLICANT: Chen, Rui-hong
 ; APPLICANT: Wang, Dunrui
 ; APPLICANT: Wang, Zhiwei

```

; APPLICANT: Wehrman, Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Qian, Xiaohong B.
; APPLICANT: Dmanac, Radoje T.
; TITLE OF INVENTION: No. US20020197679A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2
; CURRENT APPLICATION NUMBER: US/10/098,841
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 03/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 331
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 5
; LENGTH: 8062
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (363) .. (7355)
US-10-098-841-5

```

Query Match	97.7%;	Score 7614.8;	DB 13;	Length 8062;
Best Local Similarity	98.9%;	Pred. No. 0;		
Matches 7704;	Conservative 0;	Mismatches 7;	Indels 75;	Gaps 17;
2y	10	CCGCGTGTGCTGCACAGGGGAGAGAGGAAACCCAGCGCGAGCGGAAAGAGGGGACC	69	
Db	104	CCGCGTGTGCTGCACAGGGGAGAGAGGAAACCCAGCGCGAGCGGAAAGAGGGGACC	163	
2y	70	TGAGCCCAAACTTCTCTGTCTCTCTGTCAATCCCTTCTGTCTCTCTCTCAACCGCTTCCCTTCC	129	
Db	164	TGCAGCCACAACTTCTCTGTCTCTCTGTCAATCCCTTCTGTCTCTCTCAACCGCTTCCCTTCC	223	
2y	130	CACCTCTGTGGCCCCACCTTCTTGAGAGCGACAACCCCGGGAGGCATTAGAAGGGATT	189	
Db	224	CACCTCTGTGGCCCCACCTTCTTGAGAGCGACAACCCCGGGAGGCATTAGAAGGGATT	283	
Qy	190	TTCCGCGAGGTTGCGAAGGGAGCAAACTTGTTGGCAACTTGCCTTCCCGGTGCGGCGTC	249	
Db	284	TTCCGCGAGGTTGCGAAGGGAGCAAACTTGTTGGCAACTTGCCTTCCCGGTGCGGCGTC	343	
Qy	250	TCTCCCCACCGTCTCAACATGCTTTAGGGGTCCGGGCCCCGGGCTGCTGCTGGCCGT	309	
Db	344	TCTCCCCACCGTCTCAACATGCTTTAGGGGTCCGGGCCCCGGGCTGCTGCTGGCCGT	403	
Qy	310	CCTGTGCTTGCGGACAGCGGTGCCCTTCCACGGGAGCCTCGAAGAGCAAGAGCAGGCTCA	369	
Db	404	CCTGTGCTTGCGGACAGCGGTGCCCTTCCACGGGAGCCTCGAAGAGCAAGAGCAGGCTCA	463	
Qy	370	GCAATGGTTGAGCCCCAGTCCCGGTGGCTGTCACTCAAGCAAGCCCGTGTGTTATGA	429	
Db	464	GCAATGGTTGAGCCCCAGTCCCGGTGGCTGTCACTCAAGCAAGCCCGTGTGTTATGA	523	
Qy	430	CAATGGAAAACACTATCAGATAAACTCAACAGTGGGAGCGGACCTACCTAGGCAATGCGTT	489	
Db	524	CAATGGAAAACACTATCAGATAAACTCAACAGTGGGAGCGGACCTACCTAGGCAATGCGTT	583	
Qy	490	GCTTTGTACTTGTATGAGGAAGCCGAGGTTTTTAATCTCGAAAGATTAACCTTGAAGCTGA	549	
Db	584	GCTTTGTACTTGTATGAGGAAGCCGAGGTTTTTAATCTCGAGAGATAAACCCTGAAGCTGA	643	
Qy	550	AGAGACTTGTCTTTGACAAGTACACTCTGGGAACACTTACCGAGTGGGTGACACTTATGAGCG	609	
Db	644	AGAGACTTGTCTTTGACAAGTACACTCTGGGAACACTTACCGAGTGGGTGACACTTATGAGCG	703	
Qy	610	TCCTAAGAATCCCATGATCTGGACTGTACTTCGATTCGGGCTTGGCGAGGAGGAATAG	669	
Db	704	TCTTAAAGACTCAATGATCTGGACTGTACTTCGATTCGGGCTTGGCGAGGAGGAATAG	763	

Db 4004 CGGCCAGCAGGAAATCTTTGGAAGAAGTGGTCCATGCTGATCAGAGCTCCTGCACCTTT 4063
 QY 3970 TGATAAACCCTGAGTCCCGGCTCGGAGTACAATATGTCAGTGTGTTTACACTGTCTCAAGGATGACAA 4029
 Db 4064 TGAATAACCTGAGTCCCGGCTCGGAGTACAATATGTCAGTGTGTTTACACTGTCTCAAGGATGACAA 4123
 QY 4030 GGAAGTGTCCCTATCTCTGTATGATCAGATCATCCAGCTGTGTTCTCTCTCCACTGACCTGGC 4089
 Db 4124 GGAAGTGTCCCTATCTCTGTATGATCAGATCATCCAGCTGTGTTCTCTCTCCACTGACCTGGC 4183
 QY 4090 ATTCAACCAATTTGGTCCAGACACCATCGGTGTCACTGGGCTCCACCCGCCATCCATTGA 4149
 Db 4184 ATTCAACCAATTTGGTCCAGACACCATCGGTGTCACTGGGCTCCACCCGCCATCCATTGA 4243
 QY 4150 TTTAAACCAATTTCTGTGTGGTGTACTACCTGTGGAATAAGGAAAGATGTTGAGAGTT 4209
 Db 4244 TTTAAACCAATTTCTGTGTGGTGTACTACCTGTGGAATAAGGAAAGATGTTGAGAGTT 4303
 QY 4210 GTCAATTTCTCTTCAGACAATGAGTGTCTTAAACAATCTCTGCTGTGTACAGAATA 4269
 Db 4304 GTCAATTTCTCTTCAGACAATGAGTGTCTTAAACAATCTCTGCTGTGTACAGAATA 4363
 QY 4270 TGTAGTGTGTCTCCAGTGTCTAGACAACATGAGAGACACCTCTTAGAGGAAGACA 4329
 Db 4364 TGTAGTGTGTCTCCAGTGTCTAGACAACATGAGAGACACCTCTTAGAGGAAGACA 4423
 QY 4330 GAABAACAGTCTTGATTTCCCAACTGGCATTTGACTTTTCTGATATTAAGTCCCACTCTTT 4389
 Db 4424 GAABAACAGTCTTGATTTCCCAACTGGCATTTGACTTTTCTGATATTAAGTCCCACTCTTT 4483
 QY 4390 TACTGTGCACTGGATGTTGCTCTGAGCCACCATCACTGGCTACAGGATCCGCCATCATCC 4449
 Db 4484 TACTGTGCACTGGATGTTGCTCTGAGCCACCATCACTGGCTACAGGATCCGCCATCATCC 4543
 QY 4450 CGAGCACTTCAGTGGGAGACCTCGAAGAGATCGGTGCCCACTCTCGGAATTCACATCAC 4509
 Db 4544 CGAGCACTTCAGTGGGAGACCTCGAAGAGATCGGTGCCCACTCTCGGAATTCACATCAC 4603
 QY 4510 CTTCAACCACTCACTCCAGGCACAGATGATGTGGTCAGCATGTGTTCTTAATGGCAG 4569
 Db 4604 CTTCAACCACTCACTCCAGGCACAGATGATGTGGTCAGCATGTGTTCTTAATGGCAG 4663
 QY 4570 AGAGAAAGTCCCTTATTTGATTTGGCCCAACATCAACAGTTTCTGATGTTCCGAGGACCT 4629
 Db 4664 AGAGAAAGTCCCTTATTTGATTTGGCCCAACATCAACAGTTTCTGATGTTCCGAGGACCT 4723
 QY 4630 GGAAGTGTGTGTCGACCCCAACAGCCTACTGATCAGCTGGGATGCTCCTGCTGTAC 4689
 Db 4724 GGAAGTGTGTGTCGACCCCAACAGCCTACTGATCAGCTGGGATGCTCCTGCTGTAC 4783
 QY 4690 AGTGAGATATTACAGATCACTTACGGAGAACAGAGGAAATAGCCCTGTCCAGGAGTT 4749
 Db 4784 AGTGAGATATTACAGATCACTTACGGAGAACAGAGGAAATAGCCCTGTCCAGGAGTT 4843
 QY 4750 CACTGTGCTGGGAGCAAGTCTTACAGCTACCATCAGCGGCTTTAAACCTGGAGTTGATTA 4809
 Db 4844 CACTGTGCTGGGAGCAAGTCTTACAGCTACCATCAGCGGCTTTAAACCTGGAGTTGATTA 4903
 QY 4810 TACCATCACTGTGTATGCTGTCACTGGCGGTGGACAGAGCCCGCAAGCAGCCCAAT 4869
 Db 4904 TACCATCACTGTGTATGCTGTCACTGGCGGTGGACAGAGCCCGCAAGCAGCCCAAT 4963
 QY 4870 TTTCCATTAAATACCAAGAAATTTGACAAACCATCCAGATGCAAGTGAACCGATGTTCA 4929
 Db 4964 TTTCCATTAAATACCAAGAAATTTGACAAACCATCCAGATGCAAGTGAACCGATGTTCA 5023
 QY 4930 GGAACAACAGATTAGTGTCAAGTGTGCTTCAAGTGTCCCTGTACTGTTACAGAGT 4989
 Db 5024 GGAACAACAGATTAGTGTCAAGTGTGCTTCAAGTGTCCCTGTACTGTTACAGAGT 5083
 QY 4990 AACCACCACTCCCAAAAATGGACAGGACCAAAAACTAAACTGCAAGTCCAGATCA 5049
 Db 5084 AACCACCACTCCCAAAAATGGACAGGACCAAAAACTAAACTGCAAGTCCAGATCA 5143

QY 5050 AACAGAAATGACTATTGAAGGCTTCAGGCCACACAGTGGAGTATGTGGTTAGTGTCTATGC 5109
 Db 5144 AACAGAAATGACTATTGAAGGCTTCAGGCCACACAGTGGAGTATGTGGTTAGTGTCTATGC 5203
 QY 5110 TCAGAAATCCAAAGCGAGAGAGTCAAGCTCTGTGTTCAAGCTGCAAGTAAACCAATGATCG 5169
 Db 5204 TCAGAAATCCAAAGCGAGAGAGTCAAGCTCTGTGTTCAAGCTGCAAGTAAACCAATGATCG 5263
 QY 5170 CCTTAAAGGACTGGGATTCACATGATGTGGATGTGATTCATCAAAATGCTTGGGAAG 5229
 Db 5264 CCTTAAAGGACTGGGATTCACATGATGTGGATGTGATTCATCAAAATGCTTGGGAAG 5323
 QY 5230 CCCACAGGGGCAAGTTTCCAGGTACAGGTGACCTTCTCGAGCCCTGAGGATGGAATCCA 5289
 Db 5324 CCCACAGGGGCAAGTTTCCAGGTACAGGTGACCTTCTCGAGCCCTGAGGATGGAATCCA 5383
 QY 5290 TGAAGTATTTCCCTGCACTGATGTGGTGAAGAAGACATGTCAGAGCTGCAAGGCTCAGACC 5349
 Db 5384 TGAAGTATTTCCCTGCACTGATGTGGTGAAGAAGACATGTCAGAGCTGCAAGGCTCAGACC 5443
 QY 5350 GGGTTCAGAGTACACAGTCAAGTGTGGTGGCTTGCAGATGATATGGAGAGCACGCCCT 5409
 Db 5444 GGGTTCAGAGTACACAGTCAAGTGTGGTGGCTTGCAGATGATATGGAGAGCACGCCCT 5503
 QY 5410 GATTGGAACCCAGTCCACAGCTATTCTCTGACCAACTGACCTGAAAGTTCACCTCAGGTCAAC 5469
 Db 5504 GATTGGAACCCAGTCCACAGCTATTCTCTGACCAACTGACCTGAAAGTTCACCTCAGGTCAAC 5563
 QY 5470 ACCCAAGCCCTGAGCGCCAGTGGACACACCACTGTTACAGTCTCAGTGGATATCGAGT 5529
 Db 5564 ACCCAAGCCCTGAGCGCCAGTGGACACACCACTGTTACAGTCTCAGTGGATATCGAGT 5623
 QY 5530 GCGGGTGAACCCCAAGGAGAAGACCGGACCAATGAAGAAATCAACCTTGCCTGACAG 5589
 Db 5624 GCGGGTGAACCCCAAGGAGAAGACCGGACCAATGAAGAAATCAACCTTGCCTGACAG 5683
 QY 5590 CTTCATCGGTGGTGTATCAGGACTTATGTTGGTGGCCCAACCAATGAAGTGGTGTATGC 5649
 Db 5684 CTTCATCGGTGGTGTATCAGGACTTATGTTGGTGGCCCAACCAATGAAGTGGTGTATGC 5743
 QY 5650 TCTTAAAGGACACTTTTGAACAGCAGACAGCTCAGGAGTGTGACCACTCTCGAGAAATGT 5709
 Db 5744 TCTTAAAGGACACTTTTGAACAGCAGACAGCTCAGGAGTGTGACCACTCTCGAGAAATGT 5803
 QY 5710 CAGCCCAACAGAGAGGCTGTTGACAGATGCTACTGAGACCAACCATCACCATTAGCTG 5769
 Db 5804 CAGCCCAACAGAGAGGCTGTTGACAGATGCTACTGAGACCAACCATCACCATTAGCTG 5863
 QY 5770 GAGAAACCAAGACTGAGACGATCACTGGCCTCCAAAGTGTGATGCGGTTCCAGCCAAATGGCCA 5829
 Db 5864 GAGAAACCAAGACTGAGACGATCACTGGCCTCCAAAGTGTGATGCGGTTCCAGCCAAATGGCCA 5923
 QY 5830 GATCTCAATTCAGAGAAACCATCAAGCCAGATGTGAGAGCTACACCATCCTGGCTTACA 5889
 Db 5924 GATCTCAATTCAGAGAAACCATCAAGCCAGATGTGAGAGCTACACCATCCTGGCTTACA 5983
 QY 5890 ACCAGGCACTGACTCAAGATCTACCTGTACACCTTGAATGACAAATGCTCGGAGCTCCCC 5949
 Db 5984 ACCAGGCACTGACTCAAGATCTACCTGTACACCTTGAATGACAAATGCTCGGAGCTCCCC 6043
 QY 5950 TGTGGTCACTGAGCGCTCCACTGCGCATTTGATGACCAATCCAACTGCGGTTTCTGCGCAC 6009
 Db 6044 TGTGGTCACTGAGCGCTCCACTGCGCATTTGATGACCAATCCAACTGCGGTTTCTGCGCAC 6103
 QY 6010 CACACCAATTTCTTCTGCTGATCATGGCAGCGCCACCTGCGAGGATTAACCGGCTTACAT 6069
 Db 6104 CACACCAATTTCTTCTGCTGATCATGGCAGCGCCACCTGCGAGGATTAACCGGCTTACAT 6163
 QY 6070 CATCAAGTATGAGAGCTGGGTCTCCTCCAGAGAAAGTGGTCCCTCGGCCCGCCCTGG 6129
 Db 6164 CATCAAGTATGAGAGCTGGGTCTCCTCCAGAGAAAGTGGTCCCTCGGCCCGCCCTGG 6223

6130 TGTCAAGAGGCTACTATTACTGGCTGGAAACCGGAAACCGGAATATACAATTTATGTCAAT 6189
6224 TGTCAAGAGGCTACTATTCTGGCTGGAAACCGGAAACCGGAATATACAATTTATGTCAAT 6283
6190 TGGCCCTGAAGATATATCAAGAGGAGCGCCCTGATTTGGAAGGAAGAAAGACAGACGAGCT 6249
6284 TGGCCCTGAAGATATATCAAGAGGAGCGCCCTGATTTGGAAGGAAGAAAG----- 6332
6250 TCCCCAACTGGTAACCCCTTCCACACCCCAATCTTCATGGACAGAGATCTTTGGATGTTTCC 6309
6333 ----- 6332
6310 TTCCACAGTTCAAAAGACCCCTTTCTGTACCCACCCCTGGGTATGACACTGGAAATGGTAT 6369
6333 ----ACAGTTCAAAAGACCCCTTTCTGTACCCACCCCTGGGTATGACACTGGAAATGGTAT 6388
6370 TCAGCTTCTCTGGCACCTTCTGGTCAGCAACCCCAAGTGTGGGCAACCAATGATCTTTGAGGA 6429
6389 TCAGCTTCTCTGGCACCTTCTGGTCAGCAACCCCAAGTGTGGGCAACCAATGATCTTTGAGGA 6448
6430 ACATGTTTTAGGCGACACACACCGCCCAAGCGCCACCCCAATAGGCAATAGGCAAG 6489
6449 ACATGTTTTAGGCGACACACACCGCCCAAGCGCCACCCCAATAGGCAATAGGCAAG 6508
6490 ACCATACCCCGCCCAATGTAGGCAACCAAGCTCTCTCAGACCAACCATCTCATGCGCCCCC 6549
6509 ACCATACCCCGCCCAATGTAGGCAACCAAGCTCTCTCAGACCAACCATCTCATGCGCCCCC 6568
6550 ATTCAGGACACTTCTGAGTACATCATTTTCATGTCATCTCTGTTGGCACTGATGAAGAAC 6609
6569 ATTCAGGACACTTCTGAGTACATCATTTTCATGTCATCTCTGTTGGCACTGATGAAGAAC 6628
6610 CTTACAGTTTCAGGGTTCCTGGAACTTCTACCAAGTGCCTCTGACAGGCTCACCAGAGG 6669
6629 CTTACAGTTTCAGGGTTCCTGGAACTTCTACCAAGTGCCTCTGACAGGCTCACCAGAGG 6688
6670 TGCACCTCAACATCATATAGTGAGGCACTGAAGACCAAGAGGCAATAGGTTGCGGA 6729
6689 TGCACCTCAACATCATATAGTGAGGCACTGAAGACCAAGAGGCAATAGGTTGCGGA 6748
6730 AGAGTTGTTTACCGTGGGCACTCTCTCAACAGGCTTGAACCAACCTACCGATGACTC 6789
6749 AGAGTTGTTTACCGTGGGCACTCTCTCAACAGGCTTGAACCAACCTACCGATGACTC 6808
6790 GTGCTTTGACCCCTACACAGTGTCCCATATGTCGCTTGGAGATGAGTGGGAACGAATGTC 6849
6809 GTGCTTTGACCCCTACACAGTGTCCCATATGTCGCTTGGAGATGAGTGGGAACGAATGTC 6868
6850 TGAATCAGGCTTTAAACTGTTGTCAGTGTCTTGGAGTGTGATGCTTTCAGATG 6909
6869 TGAATCAGGCTTTAAACTGTTGTCAGTGTCTTGGAGTGTGATGCTTTCAGATG 6928
6910 TGATTCATCTAGATGTTGTCATCAATGTTGTAATCAAGATTTGGAGAGAGTGGGA 6969
6929 TGATTCATCTAGATGTTGTCATCAATGTTGTAATCAAGATTTGGAGAGAGTGGGA 6988
6970 CCGTCAGGAGAAATGCGCAGATGATGAGTGCACATGCTTGGAAACGGAAGAGGA 7029
6989 CCGTCAGGAGAAATGCGCAGATGATGAGTGCACATGCTTGGAAACGGAAGAGGA 7048
7030 ATTCAAGTGTGACCCCTCATGAGCAACGTTTACGATGATGGGAAGACATACCAAGTAGG 7089
7049 ATTCAAGTGTGACCCCTCATGAGCAACGTTTACGATGATGGGAAGACATACCAAGTAGG 7108
7090 AGAACAGTGCAGAGGAATATCTGGTGCAATTTGCTCTGCACATGCTTTGGAGGCA 7149
7109 AGAACAGTGCAGAGGAATATCTGGTGCAATTTGCTCTGCACATGCTTTGGAGGCA 7168
7150 GCGGGGCTGGCTGTGACAACTGCCGACACCTGGGGGTGAACCCAGTCCCGAAGGCAC 7209
7169 GCGGGGCTGGCTGTGACAACTGCCGACACCTGGGGGTGAACCCAGTCCCGAAGGCAC 7228
7210 TACTGGCCAGTCTCTACCAACCAAGTATTTCTCAGATATACCATCAGAGAACCAACTAATGT 7269

Db 7229 TACTGGCCAGTCTCTACCAACAGGATTTCTCAGAGATACCATCAGAGAACCAACTAATGT 7288
Qy 7270 TAATTTGCCCAATTAAGTGTCTTTCATGCTTTAGATGTACAGGCTGACAGAGAAGATTCGG 7329
Db 7289 TAATTTGCCCAATTAAGTGTCTTTCATGCTTTAGATGTACAGGCTGACAGAGAAGATTCGG 7348
Qy 7330 AGAGTAATATCATCTTTTCCAAATCCAGAGGAACAGAGATGTCTCTGCGCAAGATCCATCTA 7389
Db 7349 AGAGTAATATCATCTTTTCCAAATCCAGAGGAACAGAGATGTCTCTGCGCAAGATCCATCTA 7408
Qy 7390 AACTGGAGTGAATTAAGCAGACCCAGCTTAGAGTCTTCTTTCTTTTAAAGCCCTTTGC 7449
Db 7409 AACTGGAGTGAATTAAGCAGACCCAGCTTAGAGTCTTCTTTCTTTTAAAGCCCTTTGC 7468
Qy 7450 TCTGGAGGAAGTCTTCCAGCTTCAGCTCAACTCAGAGCTTCTCCAGAGCATACCCTGGGA 7509
Db 7469 TCTGGAGGAAGTCTTCCAGCTTCAGCTCAACTCAGAGCTTCTCCAGAGCATACCCTGGGA 7528
Qy 7510 GTTTCCTGAGGGTTTTCTCATAAATGAGGCTGCACATTTGCTTTGGAAGTAT 7569
Db 7529 GTTTCCTGAGGGTTTTCTCATAAATGAGGCTGCACATTTGCTTTGGAAGTAT 7588
Qy 7570 TCAATACCGTCAAGTATTTTAAATGAAGTGAATTTGAAGATTTGGATCAATAGGA 7629
Db 7589 TCAATACCGTCAAGTATTTTAAATGAAGTGAATTTGAAGATTTGGATCAATAGGA 7648
Qy 7630 AAGCATATGAGCAACCAAGATGCAAAATGTTTGAATGATATGATACCAAAATTTAAAGT 7689
Db 7649 AAGCATATGAGCAACCAAGATGCAAAATGTTTGAATGATATGATACCAAAATTTAAAGT 7708
Qy 7690 AGGAAAGTCAACCAACACTTCTGCTTTCACTTAAGTGTCTGCGCAAGTACTGTAGGA 7749
Db 7709 AGGAAAGTCAACCAACACTTCTGCTTTCACTTAAGTGTCTGCGCAAGTACTGTAGGA 7768
Qy 7750 ACAGCATGATCTTGTACTGTATATTTTAAATATCCACAGTACT 7795
Db 7769 ACAGCATGATCTTGTACTGTATATTTTAAATATCCACAGTACT 7814

RESULT 7

US-09-964-824A-574
; Sequence 574, Application US/09964824A
; Patent No. US20020102531A1
; GENERAL INFORMATION:
; APPLICANT: Horrigan, Stephen
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; FILE REFERENCE: 689290-73
; CURRENT APPLICATION NUMBER: US/09/964,824A
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/60/236,033
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,032
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,028
; PRIOR FILING DATE: 2000-09-28
; NUMBER OF SEQ ID NOS: 583
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 574
; LENGTH: 7680
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-964-824A-574

Query Match 95.2%; Score 7423; DB 9; Length 7680;
Best Local Similarity 95.8%; Pred. No. 0;
Matches 7432; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
Qy 349 GAAGAGCAAGAGGAGGCTCAGCAAAATGGTTTCAGCCCAAGTCCCGGTGCTGTCAAGTCA 408
Db 1 GAAGAGCAAGAGGAGGCTCAGCAAAATGGTTTCAGCCCAAGTCCCGGTGCTGTCAAGTCA 60

QY 409 AAGCAAGCCCGTTGTTATGCAATGGAACACATATCAGATAAATCAACAGTGGGAGCG 468
DB 61 AAGCAAGCCCGTTGTTATGCAATGGAACACATATCAGATAAATCAACAGTGGGAGCG 120
QY 469 GACCTACCTAGGCAATCGTTGGTTTGTACTGTTATGGAAGAACCGAGGTTTAACTG 528
DB 121 GACCTACCTAGGCAATCGTTGGTTTGTACTGTTATGGAAGAACCGAGGTTTAACTG 180
QY 529 CGAAGTAAACCTGAGCTGAAGAGACTTGTGTTGACAGTACACTGGGAACACTTACCG 588
DB 181 CGAAGTAAACCTGAGCTGAAGAGACTTGTGTTGACAGTACACTGGGAACACTTACCG 240
QY 589 AGTGGGTGACACTTATGAGCGTCTTAAGACTCCATGATCTGGGACTGTATCTGCAATCGG 648
DB 241 AGTGGGTGACACTTATGAGCGTCTTAAGACTCCATGATCTGGGACTGTATCTGCAATCGG 300
QY 649 GCGTGGGCGAGGAGATAAAGCTGTACCATCGCAAAACCGCTGCCATGAAGGGGTGAGTC 708
DB 301 GCGTGGGCGAGGAGATAAAGCTGTACCATCGCAAAACCGCTGCCATGAAGGGGTGAGTC 360
QY 709 CTACAAGATTGGTGACCTCGAGGAGACCAATGAGACTGGTGGTTACATGTTAGAGTG 768
DB 361 CTACAAGATTGGTGACCTCGAGGAGACCAATGAGACTGGTGGTTACATGTTAGAGTG 420
QY 769 TGTGTGCTCTGTTAATGAAAGAGGAATGGAACCTGCAAGCCCATAGCTGGAAGTGT 828
DB 421 TGTGTGCTCTGTTAATGAAAGAGGAATGGAACCTGCAAGCCCATAGCTGGAAGTGT 480
QY 829 TGATCATGCTGCTGGACTTCTATGTTGGTGGAGAACCTGGGAGAACCTACCAAGG 888
DB 481 TGATCATGCTGCTGGACTTCTATGTTGGTGGAGAACCTGGGAGAACCTACCAAGG 540
QY 889 CTGATGATGTTAGATTGTTCTGCTGGGAGAACCTGGGAGAACCTACCAAGG 948
DB 541 CTGATGATGTTAGATTGTTCTGCTGGGAGAACCTGGGAGAACCTACCAAGG 600
QY 949 TAGAAATAGATGCAACGATCAGGACACAAGACATCTTATAGAAATGAGACACCTGAG 1008
DB 601 TAGAAATAGATGCAACGATCAGGACACAAGACATCTTATAGAAATGAGACACCTGAG 560
QY 1009 CAAGAAGGATAATCAGGAAACCTGCTCAGTGGCATCTGCAAGGCAACCGCCGAGGAGA 1068
DB 661 CAAGAAGGATAATCAGGAAACCTGCTCAGTGGCATCTGCAAGGCAACCGCCGAGGAGA 720
QY 1069 GTGGAGTGTGAGGACACCTCTGTGAGACACATCGAGCGGATCTGGCCCTTCAC 1128
DB 721 GTGGAGTGTGAGGACACCTCTGTGAGACACATCGAGCGGATCTGGCCCTTCAC 780
QY 1129 CGATGTTCTGTCAGCTGTTTACCAACCGCAGCCTCACCCCGAGCCTCTCCTATGGCCA 1188
DB 781 CGATGTTCTGTCAGCTGTTTACCAACCGCAGCCTCACCCCGAGCCTCTCCTATGGCCA 840
QY 1189 CTGTGTCAAGACAGTGTGTGTTCTACTGTGGGATGAGTGGCTGAAGACACAGG 1248
DB 841 CTGTGTCAAGACAGTGTGTGTTCTACTGTGGGATGAGTGGCTGAAGACACAGG 900
QY 1249 AAATAAGCAATGCTTTGTCAGTGTCTGGCAACCGAGTCAAGTGGCAAGAGACAGCTGT 1308
DB 901 AAATAAGCAATGCTTTGTCAGTGTCTGGCAACCGAGTCAAGTGGCAAGAGACAGCTGT 960
QY 1309 AACCCAGACTTACGTTGGCAATCAAAATGGAGAGCCATGTGTCTTCAATTCACCTACAA 1368
DB 961 AACCCAGACTTACGTTGGCAATCAAAATGGAGAGCCATGTGTCTTACCAATTCACCTACAA 1020
QY 1369 TGGCAGGAGCTTACTCTCTGACACAGAGAGGGGAGAGACGACATCTTTGGTGAG 1428
DB 1021 TGGCAGGAGCTTACTCTCTGACACAGAGAGGGGAGAGACGACATCTTTGGTGAG 1080
QY 1429 CACAACTTCGAATATGACAGGACAGAAATATCTTTCTGCAAGACCACTGTTTT 1488
DB 1081 CACAACTTCGAATATGACAGGACAGAAATATCTTTCTGCAAGACCACTGTTTT 1140
QY 1489 GGTTCAGACTCGAGGAGGAAATCCAAATGGTGCCTTGTGCCACTTCCCTTCTATACAA 1548

DB 1141 GGTTCAGACTCGAGGAGGAAATCCAAATGGTGCCTTGTGCCACTTCCCTTCTATACAA 1200
QY 1549 CAAACAATTAACATGATTGCACTTTCTGAGGCGAAGAGACAAACATGAAGTGGTGG 1608
DB 1201 CAAACAATTAACATGATTGCACTTTCTGAGGCGAAGAGACAAACATGAAGTGGTGG 1260
QY 1609 GACCAACAGAAACTATGATGCCGACAGAAATTTGGGTTCTGCCCATGGCTGCCACGA 1668
DB 1261 GACCAACAGAAACTATGATGCCGACAGAAATTTGGGTTCTGCCCATGGCTGCCACGA 1320
QY 1669 GGAATCTGCACAAACAATGAAGGGTCAATGATCCGATTTGGAGATCAGTGGGATAAGCA 1728
DB 1321 GGAATCTGCACAAACAATGAAGGGTCAATGATCCGATTTGGAGATCAGTGGGATAAGCA 1380
QY 1729 GATGACATGGGTCAATGATGAGGTGACGTGTTGGGAATGGTCTGGGGAATGAC 1788
DB 1381 GATGACATGGGTCAATGATGAGGTGACGTGTTGGGAATGGTCTGGGGAATGAC 1440
QY 1789 ATGCATTTGCTACTCGCACTTCGAGATCAGTGCATTTGTTGATGACATCACTTACAATGT 1848
DB 1441 ATGCATTTGCTACTCGCACTTCGAGATCAGTGCATTTGTTGATGACATCACTTACAATGT 1500
QY 1849 GAACGACATTCACCAAGCGTCATGAAGGGGACACATGCTGAACTGTACATGCTTGG 1908
DB 1501 GAACGACATTCACCAAGCGTCATGAAGGGGACACATGCTGAACTGTACATGCTTGG 1560
QY 1909 TCAGGGTTCGGGGCAGGTGGAAGTGTATCCCTCGAACCAATGCCAGGATTCAGAGACTGG 1968
DB 1561 TCAGGGTTCGGGGCAGGTGGAAGTGTATCCCTCGAACCAATGCCAGGATTCAGAGACTGG 1620
QY 1969 GAGCTTTTATCAAAATGGAGATTCATGGGAGAAATGTCATGCTGTCAGATACAGTGG 2028
DB 1621 GAGCTTTTATCAAAATGGAGATTCATGGGAGAAATGTCATGCTGTCAGATACAGTGG 1680
QY 2029 CTACTCTATGCGCGTGGCAATTCGGGAGTGGCAATTCGCAACCTTTACAGACCTATCCAAAG 2088
DB 1681 CTACTCTATGCGCGTGGCAATTCGGGAGTGGCAATTCGCAACCTTTACAGACCTATCCAAAG 1740
QY 2089 CTCAAGTGTCTGTGCGAAGTATTTATCACTGAGACTCCGAGTCAAGCCCACTCCACCC 2148
DB 1741 CTCAAGTGTCTGTGCGAAGTATTTATCACTGAGACTCCGAGTCAAGCCCACTCCACCC 1800
QY 2149 CATCCAGTGGAAATGACACACAGCCATCTCACAATTTCCAAAGTACATTTCTCAGTGGAGACC 2208
DB 1801 CATCCAGTGGAAATGACACACAGCCATCTCACAATTTCCAAAGTACATTTCTCAGTGGAGACC 1860
QY 2209 TAAAAATTTCTGAGGCGGTTGGAAGAGACTACCATACAGGCCACTTAAACTCTTAC 2268
DB 1861 TAAAAATTTCTGAGGCGGTTGGAAGAGACTACCATACAGGCCACTTAAACTCTTAC 1920
QY 2269 CATCAAGCGCTGAAGCCTGGTGTGTATACAGGGGCCAGCTCATCAGCATCCAGCAGTA 2328
DB 1921 CATCAAGCGCTGAAGCCTGGTGTGTATACAGGGGCCAGCTCATCAGCATCCAGCAGTA 1980
QY 2329 CGGCCACCAAGAGTGAATCGCTTTGACTTCAACCAACAGCACAGCACACCTGTGAC 2388
DB 1981 CGGCCACCAAGAGTGAATCGCTTTGACTTCAACCAACAGCACAGCACACCTGTGAC 2040
QY 2389 CAGCAACCGGTGACAGGAGAGACACTCCCTTTCTCTCTTGTGGCCACTTCTGAATC 2448
DB 2041 CAGCAACCGGTGACAGGAGAGACACTCCCTTTCTCTCTTGTGGCCACTTCTGAATC 2100
QY 2449 TGTGACCGAAATCACAAGCCAGTAGCTTTGTGTTCTCTGGGTCTCAGCTTCCGACACCGT 2508
DB 2101 TGTGACCGAAATCACAAGCCAGTAGCTTTGTGTTCTCTGGGTCTCAGCTTCCGACACCGT 2160
QY 2509 GTCCGATTCGGGTGGAATATGAGCTGAGTGGGAGGAGATGAGCCACAGTACTGGA 2568
DB 2161 GTCCGATTCGGGTGGAATATGAGCTGAGTGGGAGGAGATGAGCCACAGTACTGGA 2220
QY 2569 TCTTCCAAAGCACAGCCACTTCTGTGAACATCCCTGACCTGCTTCTGGCCGAAATACAT 2628

b 2221 TCTTCAAGCAGCAGCAGCTTCTGTGAACATCCCTGACCTGCTTCTCGGCCGAAATACAT 2280
y 2629 TGTAAATGTCTATACAGATATCTGAGATGGGAGCAGAGTGTGATCTCTGTCTACTTCA 2688
b 2281 TGTAAATGTCTATACAGATATCTGAGATGGGAGCAGAGTGTGATCTCTGTCTACTTCA 2340
y 2689 AACACAGCGCTGATGCGCCCTCTGACCGGACTGTGGACCAAGTTGATGACACCTCAAT 2748
b 2341 AACACAGCGCTGATGCGCCCTCTGACCGGACTGTGGACCAAGTTGATGACACCTCAAT 2400
y 2749 TGTGTTCCGTGAGCAGACCCAGGCTCCCATACAGGGTACAGATAGTCTATTGCCC 2808
b 2401 TGTGTTCCGTGAGCAGACCCAGGCTCCCATACAGGGTACAGATAGTCTATTGCCC 2460
y 2809 ATCAGTAGAAGGTAGCAGACAGAACTCAACCTTCTGAACTGCAAACTCCGTCACCCCT 2868
b 2461 ATCAGTAGAAGGTAGCAGACAGAACTCAACCTTCTGAACTGCAAACTCCGTCACCCCT 2520
y 2869 CAGTGACTTGCAACCTGGTTCAGTATACATCACTATCTATGCTGTGGAAGAAATCA 2928
b 2521 CAGTGACTTGCAACCTGGTTCAGTATACATCACTATCTATGCTGTGGAAGAAATCA 2580
y 2929 AGAAGTACACCTGTTGATTCACAGAAACCACTGGCACCCTGAGATACAGT 2988
b 2581 AGAAGTACACCTGTTGATTCACAGAAACCACTGGCACCCTGAGATACAGT 2640
y 2989 GCCCTCTCCAGGAGCCTGCAAGTTTGTGAAGTGAAGACGCTGAAGTCAACCATCATGTG 3048
b 2641 GCCCTCTCCAGGAGCCTGCAAGTTTGTGAAGTGAAGACGCTGAAGTCAACCATCATGTG 2700
y 3049 GACACCGCTGAGAGTGCAGTACCGCTACCGTGTGATGTGATCCCGTCAACCTGCC 3108
b 2701 GACACCGCTGAGAGTGCAGTACCGCTACCGTGTGATGTGATCCCGTCAACCTGCC 2760
y 3109 TGGCAGACCGGAGAGGCTGCCATCAGCAGAAACACCTTTGAGAACTCACCGGGCT 3168
b 2761 TGGCAGACCGGAGAGGCTGCCATCAGCAGAAACACCTTTGAGAACTCACCGGGCT 2820
y 3169 GTCCCTGGGTCACCTATTAATCAAGTCTTGCAGTGGCCATGGAGGAGGCA 3228
b 2821 GTCCCTGGGTCACCTATTAATCAAGTCTTGCAGTGGCCATGGAGGAGGCA 2880
y 3229 GCCTCTGACTGTCAACAGACAAACCTGATGCTCCCACTAACCTCCAGTTTGTCAA 3288
b 2881 GCCTCTGACTGTCAACAGACAAACCTGATGCTCCCACTAACCTCCAGTTTGTCAA 2940
y 3289 TGAACATGATTTACTGTCCTGTGTGATGGACTCCACCTCGGSCCAGATAACAGGATA 3348
b 2941 TGAACATGATTTACTGTCCTGTGTGATGGACTCCACCTCGGSCCAGATAACAGGATA 3000
y 3349 CCGACTGACCGTGGGCTTACCCGAGAGAGCAGCCAGCCAGCTACAATGTGGTCCCTC 3408
b 3001 CCGACTGACCGTGGGCTTACCCGAGAGAGCAGCCAGCCAGCTACAATGTGGTCCCTC 3060
y 3409 TGTCTCCAGTACCCACTGAGGAATCTGAGCCTGATCTGAGTACACCGTATCCCTGCT 3468
b 3061 TGTCTCCAGTACCCACTGAGGAATCTGAGCCTGATCTGAGTACACCGTATCCCTGCT 3120
y 3469 GGCCTAAAGGGCAACCAAGAGAGCCCAAGCCACTGGAGTCTTTACCACTGCGAGCC 3528
b 3121 GGCCTAAAGGGCAACCAAGAGAGCCCAAGCCACTGGAGTCTTTACCACTGCGAGCC 3180
y 3529 TGGGAGCTCTATTCCACTTAAACACCGAGGTGACTGAGACCAACATTTGATCACATG 3588
b 3181 TGGGAGCTCTATTCCACTTAAACACCGAGGTGACTGAGACCAACATTTGATCACATG 3240
y 3589 GAGCCTGCTCCAGGATTTGTTTAAAGTGGGTGTACCAAGCCAGGAGGAGGAGCC 3648
b 3241 GAGCCTGCTCCAGGATTTGTTTAAAGTGGGTGTACCAAGCCAGGAGGAGGAGCC 3300
y 3649 ACCACGAGAGTGAATTCAGACTCAGGAGCATCGTTGTGTCCGGCTTGACTCCAGGAGT 3708
b 3301 ACCACGAGAGTGAATTCAGACTCAGGAGCATCGTTGTGTCCGGCTTGACTCCAGGAGT 3360

QY 3709 AGAATAGTCTACACCATCCAGTCTGAGAGATGACAGGAAAGATGCGCAATTGT 3768
Db 3361 AGAATAGTCTACACCATCCAGTCTGAGAGATGACAGGAAAGATGCGCAATTGT 3420
QY 3769 AAACAAAGTGGTGACACCATTTCTCCACCAAACTTGCATCTGGAGGCAAAACCTGA 3828
Db 3421 AAACAAAGTGGTGACACCATTTCTCCACCAAACTTGCATCTGGAGGCAAAACCTGA 3480
QY 3829 CACTGGAGTCTCACAGTCTCCTGGGAGAGGAGCACACCCAGACATTACTTGGTTATAG 3888
Db 3481 CACTGGAGTCTCACAGTCTCCTGGGAGAGGAGCACACCCAGACATTACTTGGTTATAG 3540
QY 3889 AATTACCAACCCCTACAAACGCGCAGCAGGAAATTTCTTGGAAAGAGTGTCCATGC 3948
Db 3541 AATTACCAACCCCTACAAACGCGCAGCAGGAAATTTCTTGGAAAGAGTGTCCATGC 3600
QY 3949 TGATCAGAGTCTCTGACATTTTGTATACCTGAGTCCGCGCTGGAGTACAATGTCAAGTGT 4008
Db 3601 TGATCAGAGTCTCTGACATTTTGTATACCTGAGTCCGCGCTGGAGTACAATGTCAAGTGT 3660
QY 4009 TTACACTGTCAAGGATGACAAGGAAAGTGTCCCTATCTCTGTATACCATCATCCAGCTGT 4068
Db 3661 TTACACTGTCAAGGATGACAAGGAAAGTGTCCCTATCTCTGTATACCATCATCCAGCTGT 3720
QY 4069 TCCTTCCTCCACTGACCTGCGATTCAACCAATTTGTTCAGACACCATGCTGTCACTTG 4128
Db 3721 TCCTTCCTCCACTGACCTGCGATTCAACCAATTTGTTCAGACACCATGCTGTCACTTG 3780
QY 4129 GGCTCCACCCCTCCATTCATTTAAACCACTTCTGCTGGTGTACTCACCTGTGAAAA 4188
Db 3781 GGCTCCACCCCTCCATTCATTTAAACCACTTCTGCTGGTGTACTCACCTGTGAAAA 3840
QY 4189 TGAGAAAGATGTTGACAGATTTCAATTTCTCTCAGACAAATGCGAGTGGTCTTAACAA 4248
Db 3841 TGAGAAAGATGTTGACAGATTTCAATTTCTCTCAGACAAATGCGAGTGGTCTTAACAA 3900
QY 4249 TCTCCTCCCTGGTACAGATATGTAGTGTGTCTCAGTGTCTACGAAACATGAGAG 4308
Db 3901 TCTCCTCCCTGGTACAGATATGTAGTGTGTCTCAGTGTCTACGAAACATGAGAG 3960
QY 4309 CACACCTCTTAGAGAAAGACAGAAACAGGTCTTGAATTCCTCCCACTGGGATTTTC 4368
Db 3961 CACACCTCTTAGAGAAAGACAGAAACAGGTCTTGAATTCCTCCCACTGGGATTTTC 4020
QY 4369 TGATATTACTGCCAACTCTTTTACTGTGCACTGGATTGTCTCGAGCCACCATCACTGG 4428
Db 4021 TGATATTACTGCCAACTCTTTTACTGTGCACTGGATTGTCTCGAGCCACCATCACTGG 4080
QY 4429 CTACAGGATCGCCATCATCCGAGCACTTCACTGGAGACCTCGAGAAAGATCGGTGCC 4488
Db 4081 CTACAGGATCGCCATCATCCGAGCACTTCACTGGAGACCTCGAGAAAGATCGGTGCC 4140
QY 4489 CCACCTCTGGAAATTCATCACCTTCAACCACTCAGTCCAGGACACAGATATGTGGTCAG 4548
Db 4141 CCACCTCTGGAAATTCATCACCTTCAACCACTCAGTCCAGGACACAGATATGTGGTCAG 4200
QY 4549 CATCGTTGCTTTAATGGCAGAGGAAAGTCCCTTATGATTGGCCAAACATCAACAGT 4608
Db 4201 CATCGTTGCTTTAATGGCAGAGGAAAGTCCCTTATGATTGGCCAAACATCAACAGT 4260
QY 4609 TTCTGATGTTCCGAGGACCTGGAAAGTGTGTGCGACCCCGCCAGCCCTACTGATCAG 4668
Db 4261 TTCTGATGTTCCGAGGACCTGGAAAGTGTGTGCGACCCCGCCAGCCCTACTGATCAG 4320
QY 4669 CTGGAGTGTCTCTCTGTCAAGTATTAACAGGATCACTTACGAGAAACAGAGG 4728
Db 4321 CTGGAGTGTCTCTCTGTCAAGTATTAACAGGATCACTTACGAGAAACAGAGG 4380
QY 4729 AAATAGCCCTGTCCAGAGTTCATGTGCTGGAGCAAGTCTACAGTACCATCAAGCG 4788
Db 4381 AAATAGCCCTGTCCAGAGTTCATGTGCTGGAGCAAGTCTACAGTACCATCAAGCG 4440

4789 CTTTAAACCTGGAGTTGATTATATACCATCACTGTGTATGCTGTCACTGGCGGTGGAGACAG 4848
 Db CTTTAAACCTGGAGTTGATTATATACCATCACTGTGTATGCTGTCACTGGCGGTGGAGACAG 4500
 4849 CCCGCAAGCAGCAAGCAATTTCCATTAATATACCGAACAGAAATTGACAAAACCATCCCA 4908
 Db CCCGCAAGCAGCAAGCAATTTCCATTAATATACCGAACAGAAATTGACAAAACCATCCCA 4560
 4909 GATCGAAGTGCAGATGTTGAGGACACAGCATTAAGTGTCAAGTGGTGGCTTCAGTTTC 4968
 Db GATCGAAGTGCAGATGTTGAGGACACAGCATTAAGTGTCAAGTGGTGGCTTCAGTTTC 4620
 4969 CCCTGTTACTGTTTACAGAGTAACCAACCTCCCAAAATGAGACCAAGCAACAAAAC 5028
 Db CCCTGTTACTGTTTACAGAGTAACCAACCTCCCAAAATGAGACCAAGCAACAAAAC 4680
 5029 TAAACCTGCAGGTCAGATCAACAGAAATGACTATTGAAGGCTTGACGCCACAGTGA 5088
 Db TAAACCTGCAGGTCAGATCAACAGAAATGACTATTGAAGGCTTGACGCCACAGTGA 4740
 5089 GTATGTGGTGTAGTGTCTATGCTCAGAAATCCAAAGGAGAGAGTCAGGCTCTGGTTCAGAC 5148
 Db GTATGTGGTGTAGTGTCTATGCTCAGAAATCCAAAGGAGAGAGTCAGGCTCTGGTTCAGAC 4800
 5149 TGCAGTAACCAACATTTGATGCTCCCTAAAGACTGGCATTTCACTGATGTGATGTGATTC 5208
 Db TGCAGTAACCAACATTTGATGCTCCCTAAAGACTGGCATTTCACTGATGTGATGTGATTC 4860
 5209 CATCAAAATGCTTGGGAAAGCCACAGGGGCAAGTTTCCAGGTACAGGGTGACCTACTC 5268
 Db CATCAAAATGCTTGGGAAAGCCACAGGGGCAAGTTTCCAGGTACAGGGTGACCTACTC 4920
 5269 GAGCCCTGAGGATGGAATCCATGAGCTATTCCCTGCACTGATGTTGATGATGATGATTC 5328
 Db GAGCCCTGAGGATGGAATCCATGAGCTATTCCCTGCACTGATGTTGATGATGATGATTC 4980
 5329 AGAGCTCAGAGGCTCAGACGGGTTCTGAGTACACAGTCAAGTGTGGTGGCTGGACGA 5388
 Db AGAGCTCAGAGGCTCAGACGGGTTCTGAGTACACAGTCAAGTGTGGTGGCTGGACGA 5040
 5389 TGAATGAGAGCCAGCCCTGATTGAAACCCAGTCCACAGCTATTCTCTGCAACCACTGA 5448
 Db TGAATGAGAGCCAGCCCTGATTGAAACCCAGTCCACAGCTATTCTCTGCAACCACTGA 5100
 5449 CCTGAAGTTCACTCAGTTCACACCAAGGCTGAGCGCCAGTGGACACACCCATGT 5508
 Db CCTGAAGTTCACTCAGTTCACACCAAGGCTGAGCGCCAGTGGACACACCCATGT 5160
 5509 TCAGCTCACTGGATATCGAGTGCAGGTTGACCCCAAGGAGAGACCGGACCAATGAAAGA 5568
 Db TCAGCTCACTGGATATCGAGTGCAGGTTGACCCCAAGGAGAGACCGGACCAATGAAAGA 5220
 5569 AATCAACCTTGGCTCAGACGCTCATCGGTGGTGTATCAGGACTTATGTGGCCACCA 5628
 Db AATCAACCTTGGCTCAGACGCTCATCGGTGGTGTATCAGGACTTATGTGGCCACCA 5280
 5629 ATATGAAGTGTAGTGTCTATGCTTTAAGGACACATTTGACAGAGACAGCAGCTCAGGGAGT 5688
 Db ATATGAAGTGTAGTGTCTATGCTTTAAGGACACATTTGACAGAGACAGCAGCTCAGGGAGT 5340
 5689 TGTCAACCTCTGAGAAATGTCGCCCAACAGAGGGCTGTGTGACAGATGCTACTGA 5748
 Db TGTCAACCTCTGAGAAATGTCGCCCAACAGAGGGGGCTGTGTGACAGATGCTACTGA 5400
 5749 GACCACCATCACTTGTGAGAAACCAAGACTGAGACCATCACTGGCTCCCAAGTTGA 5808
 Db GACCACCATCACTTGTGAGAAACCAAGACTGAGACCATCACTGGCTCCCAAGTTGA 5460
 5809 TGCCGTTCCAGCAATGCCAGACTCCAAATCCAGAGAACCAATCAAGCCAGATGTGAGAAG 5868
 Db TGCCGTTCCAGCAATGCCAGACTCCAAATCCAGAGAACCAATCAAGCCAGATGTGAGAAG 5520
 5869 CTACACCATCACTGGCTTACAAACAGGCACTGACTACAGATCTACCTGTACACCTTGAA 5928

5521 CTACACCATCAAGTTTACAAACGAGCAGCTGACTACAGATCTACCTGTACACCTTGAA 5580
 Db CTACACCATCAAGTTTACAAACGAGCAGCTGACTACAGATCTACCTGTACACCTTGAA 5580
 5581 TGCAATGCTGGAGCTCCCTGTGGTCACTGAGCCTCACTGCCATTTGATGACCAATC 5640
 Db TGCAATGCTGGAGCTCCCTGTGGTCACTGAGCCTCACTGCCATTTGATGACCAATC 5640
 5641 CAACCTGGTTCCTGGGCAACCAACCAATTTCTTGTGTATCATGGCAGCGCCACG 6048
 Db CAACCTGGTTCCTGGGCAACCAACCAATTTCTTGTGTATCATGGCAGCGCCACG 5700
 6049 TCCAGGATTAACCGCTACATCAAGTATGAGAGCCTGGTCTCTCTCCAGAGAGT 6108
 Db TCCAGGATTAACCGCTACATCAAGTATGAGAGCCTGGTCTCTCTCCAGAGAGT 5760
 6109 GGTCCCTGGCCCGCCCTGTGTACAGAGGCTACTATTACTGGCCTGGAAACCGGGAAC 6168
 Db GGTCCCTGGCCCGCCCTGTGTACAGAGGCTACTATTACTGGCCTGGAAACCGGGAAC 5820
 6169 CGAATATACAAATTTATGTCATTGGCCTGAAAGAAATATCAGAGAGCGAGCCCTGATTGG 6228
 Db CGAATATACAAATTTATGTCATTGGCCTGAAAGAAATATCAGAGAGCGAGCCCTGATTGG 5880
 6229 AAGGAAAGACAGAGCTTCCCAACTGGTAACCTTCCACACCCCAATCTTCATGG 6288
 Db AAGGAAAGACAGAGCTTCCCAACTGGTAACCTTCCACACCCCAATCTTCATGG 5940
 6289 ACCAGAGATCTTGAATGTTTCCCTCCACAGTTTCAAAAGACCCCTTTCTGTCACCCACCTGG 6348
 Db ACCAGAGATCTTGAATGTTTCCCTCCACAGTTTCAAAAGACCCCTTTCTGTCACCCACCTGG 6000
 6349 GTATGACACTGGAATGTTTACGCTTCTGGGCACTTCTGGTACAGAACCCAGTGTGG 6408
 Db GTATGACACTGGAATGTTTACGCTTCTGGGCACTTCTGGTACAGAACCCAGTGTGG 6060
 6409 GCAACAAATGATCTTTGAGGAACATGTTTAGGCGGACACACCGCCCAACAAAGGCCAC 6468
 Db GCAACAAATGATCTTTGAGGAACATGTTTAGGCGGACACACCGCCCAACAAAGGCCAC 6120
 6469 CCCCATAAGGCAATAGGCAACCAATACCCCGGCAATGTAGGACAAAGAGCTCTCTCA 6528
 Db CCCCATAAGGCAATAGGCAACCAATACCCCGGCAATGTAGGACAAAGAGCTCTCTCA 6180
 6529 GACAAACCTCTCATGGGCCCCATTCAGGACACTTCTGAGTACATCTTCTCATGTCATCC 6588
 Db GACAAACCTCTCATGGGCCCCATTCAGGACACTTCTGAGTACATCTTCTCATGTCATCC 6240
 6181 TGTGGCACTGTATGAGAACCCCTTACAGTTTACAGTTTCTGGAACTTCTACAGTGGCCAC 6648
 Db TGTGGCACTGTATGAGAACCCCTTACAGTTTACAGTTTCTGGAACTTCTACAGTGGCCAC 6300
 6649 TCTGACAGGCTCACCAGAGGTGCCACCTCAACATCATAGTGGAGGCACTGAAAGACCA 6708
 Db TCTGACAGGCTCACCAGAGGTGCCACCTCAACATCATAGTGGAGGCACTGAAAGACCA 6360
 6301 TCTGACAGGCTCACCAGAGGTGCCACCTCAACATCATAGTGGAGGCACTGAAAGACCA 6768
 Db TCTGACAGGCTCACCAGAGGTGCCACCTCAACATCATAGTGGAGGCACTGAAAGACCA 6420
 6769 GACAAACCTTACAGGTTTGGGAAAGAGTTTGTACCGTGGGCAACTCTGTCAACAGGCTT 6828
 Db GACAAACCTTACAGGTTTGGGAAAGAGTTTGTACCGTGGGCAACTCTGTCAACAGGCTT 6480
 6829 AGATGAGTGGGAAACCAATGCTGAAATCAGGCTTTAAATCTTGTGCCAGTCTTAGGCTT 6888
 Db AGATGAGTGGGAAACCAATGCTGAAATCAGGCTTTAAATCTTGTGCCAGTCTTAGGCTT 6540
 6889 TGGAGTGTGCTTTCAGATGTTTCTAGATGTTGAGTGGTCCATGACATGTTGTGAACCTA 6948
 Db TGGAGTGTGCTTTCAGATGTTTCTAGATGTTGAGTGGTCCATGACATGTTGTGAACCTA 6600
 6949 CAAGATTTGAGAGAGAGTGGGCGGCTCAGGAGAAATGGCCAGATGATGAGTGCACATG 7008

```
6601 CAAGATTGGAGAGAGTGGGACGTCAGGGAGAAATGCCAGATGATGAGCTGCACATG 6660
7009 TCTTGGGAACGGAAGGAGAAATTCAGTGTGACCCCTCATGAGGCAACGCTGTACGATGA 7068
6661 TCTTGGGAACGGAAGGAGAAATTCAGTGTGACCCCTCATGAGGCAACGCTGTACGATGA 6720
7069 TGGGAAGACATACCAGTATGAGGAGAACAGTGGCAGAGGAATATCTCGGTGCCATTTGCTC 7128
6721 TGGGAAGACATACCAGTATGAGGAGAACAGTGGCAGAGGAATATCTCGGTGCCATTTGCTC 6780
7129 CTGCACATGCTTTGGAGGGCCAGCGGGGCTGGGCTGTGACAACTGCGCAGACCTGGGGG 7188
6781 CTGCACATGCTTTGGAGGGCCAGCGGGGCTGGGCTGTGACAACTGCGCAGACCTGGGGG 6840
7189 TGAACCCAGTCCCGAAGGACACTACTGGCCAGTCTTACACCAAGTATCTCAGAGATACCA 7248
6841 TGAACCCAGTCCCGAAGGACACTACTGGCCAGTCTTACACCAAGTATCTCAGAGATACCA 6900
7249 TCAGAGAACAAACACTAATGTTAATTGCCCAATGAGTGTCTTCATGCTTTCAGATGTACA 7308
6901 TCAGAGAACAAACACTAATGTTAATTGCCCAATGAGTGTCTTCATGCTTTCAGATGTACA 6960
7309 GGCTGCAGAGAGATTCGGAGAGTAATCATCTTTTCCAAATCCAGAGGAAACAAGCATGT 7368
6961 GGCTGCAGAGAGATTCGGAGAGTAATCATCTTTTCCAAATCCAGAGGAAACAAGCATGT 7020
7369 CTCTCTGCCAAGATCCATCTAAACTGGAGTGTGTTAGCAGACCCAGCTTAGAGTTCTTC 7428
7021 CTCTCTGCCAAGATCCATCTAAACTGGAGTGTGTTAGCAGACCCAGCTTAGAGTTCTTC 7080
7429 TTTCTTTCTTAAGCCCTTTGCTCTGGAGGAAGTTCTCCAGCTTCAGCTCAACTCAGACT 7140
7489 TCTCCAGATCACTCCGAGTCTTCTGAGGTTTCTCATAAATAGAGGCTGACATTT 7548
7141 TCTCCAGATCACTCCGAGTCTTCTGAGGTTTCTCATAAATAGAGGCTGACATTT 7200
7549 GCCTGTCTGCTTCGAGATTTCAATACCGCTCAGTATTTTAAATGAAGTATCTAAGA 7608
7201 GCCTGTCTGCTTCGAGATTTCAATACCGCTCAGTATTTTAAATGAAGTATCTAAGA 7260
7609 TTTGGTTGGATCAATAGAGAGCATATGACGCCCAACCAAGATGCAATGTTTGAAT 7668
7261 TTTGGTTGGATCAATAGAGAGCATATGACGCCCAACCAAGATGCAATGTTTGAAT 7320
7669 GATATGACCAAAATTTAAGTAGGAAGTCAACCCAAACACTTCTGCTTTCACTTAAGTGT 7728
7321 GATATGACCAAAATTTAAGTAGGAAGTCAACCCAAACACTTCTGCTTTCACTTAAGTGT 7380
7729 CTGCGCCGCAATCTGTAGGAACAGCATGATCTTGTACTGTGATATTTTAAATATCCA 7788
7381 CTGCGCCGCAATCTGTAGGAACAGCATGATCTTGTACTGTGATATTTTAAATATCCA 7440
7789 CAGTACT 7795
7441 CAGTACT 7447
```

RESULT 8
JS-10-182-936A-75
; Sequence 75, Application US/10182936A
; Publication No. US20040038860A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Kristina M.
; APPLICANT: Anisowicz, Anthony
; APPLICANT: Bhat, Bheem
; APPLICANT: Damagnez, Veronique
; APPLICANT: Robinson, John
; APPLICANT: Yaworsky, Paul
; TITLE OF INVENTION: Reagents and Method for Modulating DKK-Mediated Interactions
; FILE REFERENCE: 032796-143
; CURRENT APPLICATION NUMBER: US/10/182,936A

```
; CURRENT FILING DATE: 2002-08-02  
; PRIOR APPLICATION NUMBER: PCT/US02/15982  
; PRIOR FILING DATE: 2002-05-17  
; PRIOR APPLICATION NUMBER: US 60/291,311  
; PRIOR FILING DATE: 2001-05-17  
; PRIOR APPLICATION NUMBER: US 60/353,058  
; PRIOR FILING DATE: 2002-02-01  
; PRIOR APPLICATION NUMBER: US 60/361,293  
; PRIOR FILING DATE: 2002-03-04  
; NUMBER OF SEQ ID NOS: 216  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 75  
; LENGTH: 7680  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-182-936A-75  
  
Query Match 95.2%; Score 7423; DB 12; Length 7680;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 7432; Conservative 0; Mismatches 15; Indels 0; Gaps 0;  
  
QY 349 GAAGAGCAAGAGGAGGCTCAGCAAAATGGTTTCAGCCCGCAGTCCCGTGGCTGTCAGTCA 408  
Db 1 GAAGAGCAAGAGGAGGCTCAGCAAAATGGTTTCAGCCCGCAGTCCCGTGGCTGTCAGTCA 60  
QY 409 AAGCAAGCCCGGTTGTTATGACAAATGGAATAAACAATATCAGATAAATCAACAGTGGAGCG 468  
Db 61 AAGCAAGCCCGGTTGTTATGACAAATGGAATAAACAATATCAGATAAATCAACAGTGGAGCG 120  
QY 469 GACCTACTAGGCAATCGTGTGTTTGTCTTGTATGAGGAGGAGCCGAGGTTTAACTG 528  
Db 121 GACCTACTAGGCAATCGTGTGTTTGTCTTGTATGAGGAGGAGCCGAGGTTTAACTG 180  
QY 529 CGAAAGTAAACCTGAAGCTGAAGAGACTTCTTTGACAACTACACTGGGAACTTACCG 588  
Db 181 CGAAAGTAAACCTGAAGCTGAAGAGACTTCTTTGACAACTACACTGGGAACTTACCG 240  
QY 589 AGTGGGTGACACTTATGAGGCTCTTAAGACTCCATGATCTGGAGCTGTACCTGCATCGG 648  
Db 241 AGTGGGTGACACTTATGAGGCTCTTAAGACTCCATGATCTGGAGCTGTACCTGCATCGG 300  
QY 649 GGCTGGGCGAGGAGAAATAAGCTGTACCATCGCAACCGCTGCCATGAAGGGGTCAAGTC 708  
Db 301 GGCTGGGCGAGGAGAAATAAGCTGTACCATCGCAACCGCTGCCATGAAGGGGTCAAGTC 360  
QY 709 CTACAAGATTGGTGACACCTTGGAGGAGACCACTGAGACTGGTGGTTACATGTTAGAGTG 768  
Db 361 CTACAAGATTGGTGACACCTTGGAGGAGACCACTGAGACTGGTGGTTACATGTTAGAGTG 420  
QY 769 TGTGTGCTTGGTAAATGGAAGAGAGAAATGGAACCTGCAAGCCCATAGCTGAGAGAGTGT 828  
Db 421 TGTGTGCTTGGTAAATGGAAGAGAGAAATGGAACCTGCAAGCCCATAGCTGAGAGAGTGT 480  
QY 829 TGATCATGCTCTGGGACTTCTTATGTTGGTGGAGAAACGTTGGAGAGGCCCTACCAAGG 888  
Db 481 TGATCATGCTCTGGGACTTCTTATGTTGGTGGAGAAACGTTGGAGAGGCCCTACCAAGG 540  
QY 889 CTGGATGATGTTAGATTGTACTTGCCTGGGAGAGGAGGAGCGGACGCTACCTTGCATTC 948  
Db 541 CTGGATGATGTTAGATTGTACTTGCCTGGGAGAGGAGGAGCGGACGCTACCTTGCATTC 600  
QY 949 TAGAAATAGATGCAACCGATCAGGACACAAGAGCATCTTATAGAAATTCGAGACACTGGAG 1008  
Db 601 TAGAAATAGATGCAACCGATCAGGACACAAGAGCATCTTATAGAAATTCGAGACACTGGAG 660  
QY 1009 CAAGAAGGATAATCGAGGAAACCTGCTCCAGTGTCTGACACAGGCAACCGCGAGGAGA 1068  
Db 661 CAAGAAGGATAATCGAGGAAACCTGCTCCAGTGTCTGACACAGGCAACCGCGAGGAGA 720  
QY 1069 GTGGAAGTGTGAGAGGCAACCTCTGTGACAGACCAATCAGAGGAGATCTGGCCCTTTCAC 1128  
Db 721 GTGGAAGTGTGAGAGGCAACCTCTGTGACAGACCAATCAGAGGAGATCTGGCCCTTTCAC 780
```

1129	CGATGTTCTGTGCAGCTGTTTACAAACCGCAGCGCTCACCCCCAGCTCTCTCCCTATGGCCA	1188
781	CGATGTTCTGTGCAGCTGTTTACAAACCGCAGCGCTCACCCCCAGCTCTCTCCCTATGGCCA	840
1189	CTGTGTCA CAGACAGTGGTGTGCTACTCTGTGGGGATGCAGTGGCTGAAGACACAAGG	1248
841	CTGTGTCA CAGACAGTGGTGTGCTACTCTGTGGGGATGCAGTGGTGAAGACACAAGG	900
1249	AAATAAGCAAAATGCTTTTGCACTGTCCTGGGCACACGAGTCAGTGCACAAGACAGCTGT	1308
901	AAATAAGCAAAATGCTTTTGCACTGTCCTGGGCACACGAGTCAGTGCACAAGACAGCTGT	960
1309	AACCCAGACTTACGGTGGCACTCAAAATGGAGAGCCATGTGTCTTACCATTCACCTACAA	1368
961	AACCCAGACTTACGGTGGCACTTAAATGGAGAGCCATGTGTCTTACCATTCACCTACAA	1020
1369	TGGCAGGAGCTTCTACTCTCTGCACCA CAGAAAGGGCGACAGGACGGACATCTTTGGTGCAG	1428
1021	TGGCAGGAGCTTCTACTCTCTGCACCA CAGAAAGGGCGACAGGACGGACATCTTTGGTGCAG	1080
1429	CACAACTTCGAAATATGAGCAGACAGACAAGAAATACTCTTTCTGCACAGACCACTGTTTT	1488
1081	CACAACCTTCGAAATATGAGCAGACAGACAAGAAATACTCTTTCTGCACAGACCACTGTTTT	1140
1489	GGTTTCAGACTCGAGGAGGAAATTCCAATGTGTCTTGTGCCACTTCCCTTCCTATACAA	1548
1141	GGTTTCAGACTCAGGAGGAAATTCCAATGTGTCTTGTGCCACTTCCCTTCCTATACAA	1200
1549	CAACCAAAATTA CACTGATTTGCACTTCTGAGGGCAGAAAGACA CAAATGAAGTGGTGTGG	1608
1201	CAACCAAAATTA CACTGATTTGCACTTCTGAGGGCAGAAAGACA CAAATGAAGTGGTGTGG	1260
1609	GACCACACAGAACTATGATGTCGCCACAGAAAGTTTGGTTCTGCCCCATGCTGCCACAGA	1668
1261	GACCACACAGAACTATGATGTCGCCACAGAAAGTTTGGTTCTGCCCCATGCTGCCACAGA	1320
1669	GGAAATCTGCACAACCAATGAAGGGGTCA TGTATCCGCA TTGGAGATCAGTGGGATGAACA	1728
1321	GGAAATCTGCACAACCAATGAAGGGGTCA TGTATCCGCA TTGGAGATCAGTGGGATGAACA	1380
1729	GCATGACATGGGTCACTATGATGAGTGCACGTGTGTGGAAATGGTCTGGGGAATGGAC	1788
1381	GCATGACATGGGTCACTATGATGAGTGCACGTGTGTGGAAATGGTCTGGGGAATGGAC	1440
1789	ATGCATTGCCCTACTCGCAGCTTCGAGATCAGTGCATTTGTATGAATCATCTTACAAATGT	1848
1441	ATGCATTGCCCTACTCGCAA CTTCGAGATCAGTGCATTTGTATGAATCATCTTACAAATGT	1500
1849	GAACGACAA TTTCCA CAAAGCGTCA TGAAGAGGGGCACATGCTGAA CTGTACATGCTTCGG	1908
1501	GAACGACAA TTTCCA CAAAGCGTCA TGAAGAGGGGCACATGCTGAA CTGTACATGCTTCGG	1560
1909	TCAGGGTCTGGGGCAGGTGGAAAGTGTGATCCGCTCGACCAATGCCAGGATTCAGAGACTGG	1968
1561	TCAGGGTCTGGGGCAGGTGGAAAGTGTGATCCGCTCGACCAATGCCAGGATTCAGAGACTGG	1620
1969	GACGTTTTATCAAAATGAGAA TTTATCGGGAGAAATATGTGCATGGTGTGCAGATCCAGTG	2028
1621	GACGTTTTATCAAAATGAGAA TTTATCGGGAGAAATATGTGCATGGTGTGCAGATCCAGTG	1680
2029	CTACTGCTATGGCGTGGCA TTTGGGAGTGCATTTGCCA ACTTTACAGACCTTATCCAAG	2088
1681	CTACTGCTATGGCGTGGCA TTTGGGAGTGGCA TTTGCCA ACTTTACAGACCTTATCCAAG	1740
2089	CTCAAGTGGTCTGTTCGAAGTATTTATCACTGAGACTCCGAGTCAGCCCCAACTCCCAACC	2148
1741	CTCAAGTGGTCTGTTCGAAGTATTTATCACTGAGACTCCGAGTCAGCCCCAACTCCCAACC	1800
2149	CATCCAGTGGAA TGCACCA CAGCCACTCTACATTTCCAGTATCATTTCTCAGGTGGAGACC	2208
1801	CATCCAGTGGAA TGCACCA CAGCCACTCTACATTTCCAGTATCATTTCTCAGGTGGAGACC	1860
2209	TAAAAATTTCTGTAGCCGCTTGGAGGAAAGCTACCA TACCAGGGCCACTTAAACTCTTACAC	2268
QY		QY

1861	TAAAAATTTCTGTAGGCGGTTGGAGGAAGCTACCATACGAGGCCACTTAAACTCCTACAC	1920
2269	CATCAAAAGGCTGAAGCCTGGTGTGGTATACAGAGGGCCAGTCTATCAGCATCCACGAGTA	2328
1921	CATCAAAAGGCTGAAGCCTGGTGTGGTATACAGAGGGCCAGTCTATCAGCATCCACGAGTA	1980
2329	CGGCCACCAAGAGTACTGCTTTCAGCTTCCACACCCAGCACCAGCAGCACCTGTGAC	2388
1981	CGGCCACCAAGAGTACTGCTTTCAGCTTCCACACCCAGCACCAGCAGCACCTGTGAC	2040
2389	CAGCAACACCGTGACGAGCAGACGACTCCCTTTTCTCTCTTTGGGCCACTTCTGTAATC	2448
2041	CAGCAACACCGTGACGAGCAGACGACTCCCTTTTCTCTCTTTGGGCCACTTCTGTAATC	2100
2449	TGTGACCGAAATCAACGCATGACTTTGTGGTCTCTGGGTCTCAGCTTCGGACACCGT	2508
2101	TGTGACCGAAATCAACGCATGACTTTGTGGTCTCTGGGTCTCAGCTTCGGACACCGT	2160
2509	GTCCGGATTCCGGGTGGAAATAGCTGTGAGTGGAGGGAGATGAGCCACAGTACCTGGGA	2568
2161	GTCCGGATTCCGGGTGGAAATAGCTGTGAGTGGAGGGAGATGAGCCACAGTACCTGGGA	2220
2569	TCTTCCAAAGCACGCCACTCTCTGTGAACATCCCTGACCTGTCTCTGGCGGAAATACAT	2628
2221	TCTTCCAAAGCACGCCACTCTCTGTGAACATCCCTGACCTGTCTCTGGCGGAAATACAT	2280
2629	TGTAAATGCTTATCAGATATCTGAGATGGGGAGCAGATTTGATCTCTACTTTCACA	2688
2281	TGTAAATGCTTATCAGATATCTGAGATGGGGAGCAGATTTGATCTCTACTTTCACA	2340
2689	AACAAACAGCGCTGATGCCCTCTCTGACCCGACTGTGGACCAAGTTGATGACACCTCAAT	2748
2341	AACAAACAGCGCTGATGCCCTCTCTGACCCGACTGTGGACCAAGTTGATGACACCTCAAT	2400
2749	TGTTGTTGCTGGAGCAGACCCAGGCTCCCATCAGAGGTCAGAAATAGTCTATTGCC	2808
2401	TGTTGTTGCTGGAGCAGACCCAGGCTCCCATCAGAGGTCAGAAATAGTCTATTGCC	2460
2809	ATCAGTAGAAGTAGCAGCACAGAACTCAACCTTCTCGAAATCGCAAACTCCGTCACCT	2868
2461	ATCAGTAGAAGTAGCAGCACAGAACTCAACCTTCTCGAAATCGCAAACTCCGTCACCT	2520
2869	CAGTGACTTGCAAACTGGTGTTCAGTATACATCATCTATGCTGTGGAGAAATCA	2928
2521	CAGTGACTTGCAAACTGGTGTTCAGTATACATCATCTATGCTGTGGAGAAATCA	2580
2929	AGAAAGTACACTGTGTCTTCAACAGAGAAACCATCTGCGACCCACGCTCAGATACAGT	2988
2581	AGAAAGTACACTGTGTGTCTTCAACAGAGAAACCATCTGCGACCCACGCTCAGATACAGT	2640
2989	GCCCTCTCCAGGACCTGCAGTTTGGAGTGCACAGCTGAAGGTCAACATCATGTG	3048
2641	GCCCTCTCCAGGACCTGCAGTTTGGAGTGCACAGCTGAAGGTCAACATCATGTG	2700
3049	GACACGGCTGAGAGTGCAGTGCAGGCTACCGTGTGGATGTGATCCCGCTCACCTGCC	3108
2701	GACACGGCTGAGAGTGCAGTGCAGGCTACCGTGTGGATGTGATCCCGCTCACCTGCC	2760
3109	TGGCAGCAGCGGCAGAGGCTGCCCATCAGCAGGAACACTTTTGAGAGTCAACCGGCT	3168
2761	TGGCAGCAGCGGCAGAGGCTGCCCATCAGCAGGAACACTTTTGAGAGTCAACCGGCT	2820
3169	GTCCCTGGGGTCACTTATTACTTCAAAGTCTTTGAGTGAGCCATGGGAGGAGAGCAA	3228
2821	GTCCCTGGGGTCACTTATTACTTCAAAGTCTTTGAGTGAGCCATGGGAGGAGAGCAA	2880
3229	GCCTTGACTGCTCAACAGACAAACAACTGGATGCTCCCACTTAACCTCCAGTTTGTCAA	3288
2881	GCCTTGACTGCTCAACAGACAAACAACTGGATGCTCCCACTTAACCTCCAGTTTGTCAA	2940
3289	TGAAACTGATTCTACTGTCTGTGTGAGATGGAATCCACTCTGGGGCCAGATTAACAGGATA	3348

2941 TGAACCTGATTTCTCTGCTGAGATGAGATCCACCTCGGGCCAGATTAACAGGATA 3000
 3349 CCACCTGACCGTGGGCTTACCCGAGAGAGACAGCCAGCAGTACAAATGCGGTCCCTC 3408
 3001 CCACCTGACCGTGGGCTTACCCGAGAGAGACAGCCAGCAGTACAAATGCGGTCCCTC 3060
 3409 TGCTCTCAAGTACCCACTGAGGAATCTGAGCGCTGATCTGAGTACACCGTATCCCTGT 3468
 3061 TGCTCTCAAGTACCCACTGAGGAATCTGAGCGCTGATCTGAGTACACCGTATCCCTGT 3120
 3469 GGCCTATTAAGGGCAACCAAGAGAGCCCAAGGCACTGGAGTCTTTACCACTGCGAGCC 3528
 3121 GGCCTATTAAGGGCAACCAAGAGAGCCCAAGGCACTGGAGTCTTTACCACTGCGAGCC 3180
 3529 TGGAGCTCTATTCCACCTTACAAACCGAGTGTACTGAGACCAACATTTGATCACATG 3588
 3181 TGGAGCTCTATTCCACCTTACAAACCGAGTGTACTGAGACCAACATTTGATCACATG 3240
 3589 GAGCCTGTCCCAAGATTTGTTTAAAGCTGGGTGTACGACCAAGCCAGGAGGAGGC 3648
 3241 GAGCCTGTCTCCCAAGATTTGTTTAAAGCTGGGTGTACGACCAAGCCAGGAGGAGGC 3300
 3649 ACCACGAGAGTGTACTGAGTACAGTACAGGAGCATCGTTGTCCGGTTCACTCCAGGAGT 3708
 3301 ACCACGAGAGTGTACTGAGTACAGTACAGGAGCATCGTTGTCCGGTTCACTCCAGGAGT 3360
 3709 AGAATAGCTCTACCACTCAAGTCTGAGATGACAGGAGGAGATGCGCAATTTGT 3768
 3361 AGAATAGCTCTACCACTCAAGTCTGAGATGACAGGAGGAGATGCGCAATTTGT 3420
 3769 AAACAAAGTGTGACACCATTTGTCTCCACCAACAAACTTGCATCTGGAGGCAAAACCTGA 3828
 3421 AAACAAAGTGTGACACCATTTGTCTCCACCAACAAACTTGCATCTGGAGGCAAAACCTGA 3480
 3829 CACTGGAGTCTCAGTCTCTGGAGAGGAGACCAACCCAGACATTACTGTTATAG 3888
 3481 CACTGGAGTCTCAGTCTCTGGAGAGGAGACCAACCCAGACATTACTGTTATAG 3540
 3889 AATTACCAACCCCTACAAACCGGAGGAGGAAATTTCTTGGAGAGAGTGTCCATGC 3948
 3541 AATTACCAACCCCTACAAACCGGAGGAGGAAATTTCTTGGAGAGAGTGTCCATGC 3600
 3949 TGATCAGAGTCTCTGACCTTTGATACCTGAGTCCGGCTGAGTACAAATGTCAGTGT 4008
 3601 TGATCAGAGTCTCTGACCTTTGATACCTGAGTCCGGCTGAGTACAAATGTCAGTGT 3660
 4009 TTACACTGTCAAGATGACAGGAAGTGTCCCTATCTCTGATACCATCATCCAGCTGT 4068
 3661 TTACACTGTCAAGATGACAGGAAGTGTCCCTATCTCTGATACCATCATCCAGCTGT 3720
 4069 TCCTCTCCCACTGACCTGCAATTCACCAACATGTTGTCAGACACCATGCGTGTACCTG 4128
 3721 TCCTCTCCCACTGACCTGCAATTCACCAACATGTTGTCAGACACCATGCGTGTACCTG 3780
 4129 GGTCTCACCCCTCCATTCATTAACCAACTTCTGGTGGTGTACTCTACCTGTGAAAAA 4188
 3781 GGTCTCACCCCTCCATTCATTAACCAACTTCTGGTGGTGTACTCTACCTGTGAAAAA 3840
 4189 TGAGGAAGATGTTTCAGAGTGTCAATTTCTCTTCAAGAACATGAGTGTCTTAAACAAA 4248
 3841 TGAGGAAGATGTTTCAGAGTGTCAATTTCTCTTCAAGAACATGAGTGTCTTAAACAAA 3900
 4249 TCCTCTCCCTGTTACAGAAATGTTAGTGTGAGTGTCTCCAGTGTCTACGAACACATGAGAG 4308
 3901 TCCTCTCCCTGTTACAGAAATGTTAGTGTGAGTGTCTCCAGTGTCTACGAACACATGAGAG 3960
 4309 CACACCTCTTTAGAGGAAGACAGAAAAACAGTCTTGATTTCCCACTGGCATTTGACTTTTC 4368
 3961 CACACCTCTTTAGAGGAAGACAGAAAAACAGTCTTGATTTCCCACTGGCATTTGACTTTTC 4020
 4369 TGATATTACTGCCAATCTTTTACTGTGCACTGGATTGCTCTCGAGGCCACCATCACTG 4428
 4021 TGATATTACTGCCAATCTTTTACTGTGCACTGGATTGCTCTCGAGGCCACCATCACTG 4080

QY 4429 CTACAGGATCCGCATCATCCCGAGCACTTCACTGGAGACCTCGAGAGATCGGTGCC 4488
 Db 4081 CTACAGGATCCGCATCATCCCGAGCACTTCACTGGAGACCTCGAGAGATCGGTGCC 4140
 QY 4489 CCACCTCTCGAATTCATCACTCCCTCACCACCTCACTCCAGGACACAGATGATGTGTGTCAG 4548
 Db 4141 CCACCTCTCGAATTCATCACTCCCTCACCACCTCACTCCAGGACACAGATGATGTGTGTCAG 4200
 QY 4549 CATCGTTGCTCTTAATGGCAGAGAGAAAGTCCCTTATTGATTGGCCAAACAATCAACAGT 4608
 Db 4201 CATCGTTGCTCTTAATGGCAGAGAGAAAGTCCCTTATTGATTGGCCAAACAATCAACAGT 4260
 QY 4609 TTCTGATGTTCCGAGGACCTGGAAGTGTGTGGACCCGCCACGACCTACTGATCAG 4668
 Db 4261 TTCTGATGTTCCGAGGACCTGGAAGTGTGTGGACCCGCCACGACCTACTGATCAG 4320
 QY 4669 CTGGGATGCTCTCTGCTGTACAGTGTGATATTACAGATCACTTACGGAGAAAACAGGAGG 4728
 Db 4321 CTGGGATGCTCTCTGCTGTACAGTGTGATATTACAGATCACTTACGGAGAAAACAGGAGG 4380
 QY 4729 AAATAGCCCTGTCCAGAGTTCACGTGCTGGAGCAAGTCTACAGTACCATCAGCGG 4788
 Db 4381 AAATAGCCCTGTCCAGAGTTCACGTGCTGGAGCAAGTCTACAGTACCATCAGCGG 4440
 QY 4789 CCTTAAACCTGGAGTTCATTAATACCATCACTGTGTATGCTGTCACTGGCCGTGGAGACAG 4848
 Db 4441 CCTTAAACCTGGAGTTCATTAATACCATCACTGTGTATGCTGTCACTGGCCGTGGAGACAG 4500
 QY 4849 CCGCGAAGCAGCAGCAATTTCCATTAATACCGAACAGAAATTCACAAACCATCCCA 4908
 Db 4501 CCGCGAAGCAGCAGCAATTTCCATTAATACCGAACAGAAATTCACAAACCATCCCA 4560
 QY 4909 GATGCAAGTGTACCATGCTCAGGACCAAGATTAAGTGTCAAGTGGCTGCTTCAAGTTC 4968
 Db 4561 GATGCAAGTGTACCATGCTCAGGACCAAGATTAAGTGTCAAGTGGCTGCTTCAAGTTC 4620
 QY 4969 CCCTGTACTGTTTACAGATTAACCACTCCCAAAATGGACAGGACCAACAAAC 5028
 Db 4621 CCCTGTACTGTTTACAGATTAACCACTCCCAAAATGGACAGGACCAACAAAC 4680
 QY 5029 TAAACCTGACGCTCCAGATCAACACAGAAATGATATTGAAAGGCTTGCAGCCACAGTGA 5088
 Db 4681 TAAACCTGACGCTCCAGATCAACACAGAAATGATATTGAAAGGCTTGCAGCCACAGTGA 4740
 QY 5089 GTATGTGTGTAGTGTCTATGCTCAGAAATCCAAAGCGAGAGATCAGCCTCTGTGTTCAAG 5148
 Db 4741 GTATGTGTGTAGTGTCTATGCTCAGAAATCCAAAGCGAGAGATCAGCCTCTGTGTTCAAG 4800
 QY 5149 TGCAGTAAACCAACATTCATCGCCCTTAAGGACTGGCATTCACATGTGTGATGTCGATTC 5208
 Db 4801 TGCAGTAAACCAACATTCATCGCCCTTAAGGACTGGCATTCACATGTGTGATGTCGATTC 4860
 QY 5209 CATCAAAATTTGCTTGGGAAAGCCACAGGGGCAAGTTTCCAGGTACAGGTGACCTACTC 5268
 Db 4861 CATCAAAATTTGCTTGGGAAAGCCACAGGGGCAAGTTTCCAGGTACAGGTGACCTACTC 4920
 QY 5269 GAGCCTGTGAGATGGAATCCATGAGCTATTCCCTGACCTGATGGTGAAGAGACACTGC 5328
 Db 4921 GAGCCTGTGAGATGGAATCCATGAGCTATTCCCTGACCTGATGGTGAAGAGACACTGC 4980
 QY 5329 AGAGCTGCAAGGCTCTCAGACCGGGTCTTGAGTACACAGTCAAGTGTGTTGCTTGCACGA 5388
 Db 4981 AGAGCTGCAAGGCTCTCAGACCGGGTCTTGAGTACACAGTCAAGTGTGTTGCTTGCACGA 5040
 QY 5389 TGATATGGAGAGCCAGCCCTGATTTGGAACCCAGTCCACAGCTATTCTCTGACCACTGA 5448
 Db 5041 TGATATGGAGAGCCAGCCCTGATTTGGAACCCAGTCCACAGCTATTCTCTGACCACTGA 5100
 QY 5449 CCTGAAGTTCACCTCAGGTTCACACCCCAAGCCTGAGGCCCGCAGTGGACACCACTATGT 5508
 Db 5101 CCTGAAGTTCACCTCAGGTTCACACCCCAAGCCTGAGGCCCGCAGTGGACACCACTATGT 5160

QY 5509 TCAGCTCACTGGATATCGAGTGGGGTGACCCCAAGGAGAGACCGGACCAATGAAAGA 5568
DB 5161 TCAGCTCACTGGATATCGAGTGGGGTGACCCCAAGGAGAGACCGGACCAATGAAAGA 5220
QY 5569 AATCAACCTTCTCTCTGACAGCTCATCCGTGGTGTATACAGGACTTATGGTGGCCACCAA 5628
DB 5221 AATCAACCTTCTCTCTGACAGCTCATCCGTGGTGTATACAGGACTTATGGTGGCCACCAA 5280
QY 5629 ATATGAGTGTAGTGTCTATGCTCTTAAGGACACTTTGACAGCAGACGACTCAGGAGT 5688
DB 5281 ATATGAGTGTAGTGTCTATGCTCTTAAGGACACTTTGACAGCAGACGACTCAGGAGT 5340
QY 5689 TGTCAACACTCTGGAGAAATGTGAGCCCAACCAAGAGGCTCGTGTGACAGATGCTACTGA 5748
DB 5341 TGTCAACACTCTGGAGAAATGTGAGCCCAACCAAGAGGCTCGTGTGACAGATGCTACTGA 5400
QY 5749 GACACCATACCATTAAGTGTGGAGAACCAAGACTGAGACGATCACTGGCTCCAGTTGA 5808
DB 5401 GACCACCATACCATTAAGTGTGGAGAACCAAGACTGAGACGATCACTGGCTCCAGTTGA 5460
QY 5809 TGCGTTCCAGCCAAATGGCCAGACTCCCAATCCAGAGAACCAATCAAGCCAGATGTCAGAAG 5868
DB 5461 TGCGTTCCAGCCAAATGGCCAGACTCCCAATCCAGAGAACCAATCAAGCCAGATGTCAGAAG 5520
QY 5869 CTACACCATCACTGCTTTACCAACAGGCACTGACTAACAAGATCTACCTGTACACCTTGAA 5928
DB 5521 CTACACCATCACTGCTTTACCAACAGGCACTGACTAACAAGATCTACCTGTACACCTTGAA 5580
QY 5929 TGCAATGCTCGAGCTCCCTGCTGTGTCATCAGCGCTCCACTGCCATTTGATGACCACTC 5988
DB 5581 TGCAATGCTCGAGCTCCCTGCTGTGTCATCAGCGCTCCACTGCCATTTGATGACCACTC 5640
QY 5989 CAACCTGCGTTTCCCTGGCCACCAACCAACCAATTCCTTGTCTGTATCATGGCAGCCGCAAC 6048
DB 5641 CAACCTGCGTTTCCCTGGCCACCAACCAACCAATTCCTTGTCTGTATCATGGCAGCCGCAAC 5700
QY 6049 TGCCAGGATTAACCGTACATCATCAAGTATGAGAGCGCTGGTCTCTCTCCAGAGAAT 6108
DB 5701 TGCCAGGATTAACCGTACATCATCAAGTATGAGAGCGCTGGTCTCTCTCCAGAGAAT 5760
QY 6109 GGTCCCTCGGCCCCCTGCTGTGTCACAGAGGCTACTATTACTGGCCTGGAAACCGGCAAC 6168
DB 5761 GGTCCCTCGGCCCCCTGCTGTGTCACAGAGGCTACTATTACTGGCCTGGAAACCGGCAAC 5820
QY 6169 CGAATATACAAATTTATGTCATTCGCTCAAGAAATATCAGAGAGCGAGCCCTGATTGG 6228
DB 5821 CGAATATACAAATTTATGTCATTCGCTCAAGAAATATCAGAGAGCGAGCCCTGATTGG 5880
QY 6229 AAGGAAAAGACAGACGAGCTTCCCAACTGGTAACCTTCCACACCCCAATCTTTCATGG 6288
DB 5881 AAGGAAAAGACAGACGAGCTTCCCAACTGGTAACCTTCCACACCCCAATCTTTCATGG 5940
QY 6289 ACCAGAGATCTTGGATGTTCTTCCACAGTTCAAAGAACCCTTTCGTCACCCACCTTGG 6348
DB 5941 ACCAGAGATCTTGGATGTTCTTCCACAGTTCAAAGAACCCTTTCGTCACCCACCTTGG 6000
QY 6349 GPATCAGACTGGAATGTTATTCAGCTTCTCGGCACTTCTGTCAGCAACCCAGTGTGG 6408
DB 6001 GPATCAGACTGGAATGTTATTCAGCTTCTCGGCACTTCTGTCAGCAACCCAGTGTGG 6060
QY 6409 GCAACAAATGATCTTTGAGGAACATGGTTTTAGGCGGACCAACCGGCCCAACAAAGCGGCAC 6468
DB 6061 GCAACAAATGATCTTTGAGGAACATGGTTTTAGGCGGACCAACCGGCCCAACAAAGCGGCAC 6120
QY 6469 CCCCATAAGGCATAGGCCAAGACCATACCCCGGAATGTAAGGACCAAGAGCTCTCTCTCA 6528
DB 6121 CCCCATAAGGCATAGGCCAAGACCATACCCCGGAATGTAAGGACCAAGAGCTCTCTCTCA 6180
QY 6529 GACAAACCATCTCATGGGCCCAATTCAGGACACTTCTGAGTACATCAATTTATGTCATCC 6588
DB 6181 GACAAACCATCTCATGGGCCCAATTCAGGACACTTCTGAGTACATCAATTTATGTCATCC 6240
QY 6589 TGTGGCACTGATGAGAACCCCTTACAGTTACAGTTACAGGTTCTTGGAACTTCTACCAAGTGCAC 6648

DB 6241 TGTGGCACTGATGAGAACCCCTTACAGTTACAGGTTCTCTGGAACCTTCTACCAAGTGCAC 6300
QY 6649 TCTGACAGGCTTCAACAGAGTGGCCACCTCAACAATCATAGTGGAGGACCTGAAGAACA 6708
DB 6301 TCTGACAGGCTTCAACAGAGTGGCCACCTCAACAATCATAGTGGAGGACCTGAAGAACA 6360
QY 6709 GCAGAGGCATAAGGTTCCGGGAAGAGGTTGTACCGTGGGCAACTCTGTCAACAGAGGCTT 6768
DB 6361 GCAGAGGCATAAGGTTCCGGGAAGAGGTTGTACCGTGGGCAACTCTGTCAACAGAGGCTT 6420
QY 6769 GAAACCAACCTACGAGTACTCGTGTGTGACCCCTACACAGTGTCCCAATTTATGCCGTTGG 6828
DB 6421 GAAACCAACCTACGAGTACTCGTGTGTGACCCCTACACAGTGTCCCAATTTATGCCGTTGG 6480
QY 6829 AGATGAGTGGGAACCAATGTCTGAATCAGGCTTTAAACTGTGTGCCAGTGTCTAGGCTT 6888
DB 6481 AGATGAGTGGGAACCAATGTCTGAATCAGGCTTTAAACTGTGTGCCAGTGTCTAGGCTT 6540
QY 6889 TGGAAATGCTCAATTTTCAGATGTATTCATCTAGATGGTGCACATGAATGGTGTGAACCTA 6948
DB 6541 TGGAAATGCTCAATTTTCAGATGTATTCATCTAGATGGTGCACATGAATGGTGTGAACCTA 6600
QY 6949 CAAGATTGGAGAGAGTGGGACCGTCAAGGAGAAATGGCCAGATGATGAGCTGCACATG 7008
DB 6601 CAAGATTGGAGAGAGTGGGACCGTCAAGGAGAAATGGCCAGATGATGAGCTGCACATG 6660
QY 7009 TCTTGGGAACGGAAGAGAAATTCAGATGTCAACCTCATGAGGCAACGCTGTTCAGATGA 7068
DB 6661 TCTTGGGAACGGAAGAGAAATTCAGATGTCAACCTCATGAGGCAACGCTGTTCAGATGA 6720
QY 7069 TGGGAAGACATACCACTAGGAGAACAGTGGCAGAGAAATCTCGGTGCCATTTGCTC 7128
DB 6721 TGGGAAGACATACCACTAGGAGAACAGTGGCAGAGAAATCTCGGTGCCATTTGCTC 6780
QY 7129 CTGCACATCTTTTGGAGGCGAGCGGCTGGCGCTGTGACAACTGCCGAGACCTGGGG 7188
DB 5781 CTGCACATCTTTTGGAGGCGAGCGGCTGGCGCTGTGACAACTGCCGAGACCTGGGG 6840
QY 7189 TGAAACCACTCCCGAAGGCACTACTGGCCAGTCTCAACACAGTATTTCTCAGAGATACCA 7248
DB 6841 TGAAACCACTCCCGAAGGCACTACTGGCCAGTCTCAACACAGTATTTCTCAGAGATACCA 6900
QY 7249 TCAGAGAAACAAACACTTAATGTTAAATGGCCCAATTCAGTGTCTTCAATGCCCTTTAGATGTACA 7308
DB 6901 TCAGAGAAACAAACACTTAATGTTAAATGGCCCAATTCAGTGTCTTCAATGCCCTTTAGATGTACA 6960
QY 7309 GGCTGACAGAGAAGATTCGCCGAGAGTAAATCATCTTTCCAAATCCAGAGAAACAAGCATGT 7368
DB 6961 GGCTGACAGAGAAGATTCGCCGAGAGTAAATCATCTTTCCAAATCCAGAGAAACAAGCATGT 7020
QY 7369 CTCTCTGCCAAGATCCATCTAACTGGAGTGTGTAGCAGACCCAGCTTAGAGTTCTTC 7428
DB 7021 CTCTCTGCCAAGATCCATCTAACTGGAGTGTGTAGCAGACCCAGCTTAGAGTTCTTC 7080
QY 7429 TTTCTTTTCTTAAGCCCTTTGCTCTGGAGGAAGTTCTTCCAGCTTTCAGCTCAACTCAGCT 7488
DB 7081 TTTCTTTTCTTAAGCCCTTTGCTCTGGAGGAAGTTCTTCCAGCTTTCAGCTCAACTCAGCT 7140
QY 7489 TCTCCAGCATCACCCTGGAGTTTCTGAGGTTTCTCATATAATGAGGCTGCACATT 7548
DB 7141 TCTCCAGCATCACCCTGGAGTTTCTGAGGTTTCTCATATAATGAGGCTGCACATT 7200
QY 7549 GCCTGTTCTGCTTCCGAAGTATTCAAATACCGCTCAGTATTTTAAATGAAGTGTCTTAAGA 7608
DB 7201 GCCTGTTCTGCTTCCGAAGTATTCAAATACCGCTCAGTATTTTAAATGAAGTGTCTTAAGA 7260
QY 7609 TTTGGTTGGATCAATAGGAAGCATATGCGGCCCAACCAAGATGCAATGTTTGAAT 7668
DB 7261 TTTGGTTGGATCAATAGGAAGCATATGCGGCCCAACCAAGATGCAATGTTTGAAT 7320
QY 7669 GATATGACCAAAATTTAAATAGGAAGTCAACCCAAACACTTCTGCTTTCACTTAAAGTGT 7728

b 7321 GATATGACCAAAATTTTAAGTAGGAAAGTCACCCAAAACACTTCTGCTTTCACTTAAGTGT 7380
 Y 7729 CTGGCCCGCAATACCTGTAGGAAACAGCATGATCTTGTACTGTGATATTTTAAATATCCA 7788
 b 7381 CTGGCCCGCAATACCTGTAGGAAACAGCATGATCTTGTACTGTGATATTTTAAATATCCA 7440
 Y 7789 CAGTACT 7795
 b 7441 CAGTACT 7447

RESULT 9
 S-10-171-311-63
 Sequence 63, Application US/10171311
 Publication No. US20030087270A1
 GENERAL INFORMATION:
 APPLICANT: Chen, Robert
 APPLICANT: Schlegel, Robert
 APPLICANT: Chen, Yan
 APPLICANT: Zhao, Xumei
 APPLICANT: Monahan, John
 APPLICANT: Kamatkar, Shubhangi
 APPLICANT: Glatt, Karen
 APPLICANT: Gannavarapu, Manjula
 APPLICANT: Hoerish, Sebastian
 TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
 TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
 TITLE OF INVENTION: OF CERVICAL CANCER
 FILE REFERENCE: MRI-035
 CURRENT APPLICATION NUMBER: US/10/171,311
 CURRENT FILING DATE: 2002-06-12
 PRIOR APPLICATION NUMBER: US 60/298,159
 PRIOR FILING DATE: 2001-06-13
 PRIOR APPLICATION NUMBER: US 60/298,155
 PRIOR FILING DATE: 2001-06-13
 PRIOR APPLICATION NUMBER: US 60/335,936
 PRIOR FILING DATE: 2001-11-14
 NUMBER OF SEQ ID NOS: 238
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 63
 LENGTH: 7680
 TYPE: DNA
 ORGANISM: Homo sapiens
 S-10-171-311-63

Query Match 95.2%; Score 7423; DB 14; Length 7680;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 7432; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Y 349 GAAGAGCAAGAGGAGGCTCAGCAAAATGGTTACGCCCAAGTCCCGGTGGCTGTGAGTCA 408
 b 1 GAAGAGCAAGAGGAGGCTCAGCAAAATGGTTACGCCCAAGTCCCGGTGGCTGTGAGTCA 60
 Y 409 AAGCAAGCCGGTTGTATGACAAATGAGAAACACTATCATGATAATCAACAGTGGAGCG 468
 b 61 AAGCAAGCCGGTTGTATGACAAATGAGAAACACTATCATGATAATCAACAGTGGAGCG 120
 Y 469 GACCTACCTAGGCAAGCGTGTGTTGTACTTCTTATGAGGAAGCCGAGGTTTAACTG 528
 b 121 GACCTACCTAGGCAAGCGTGTGTTGTACTTCTTATGAGGAAGCCGAGGTTTAACTG 180
 Y 529 CGAAAGTAAACCTGAAAGTGAAGAGACTTGTCTTTGACAACTGACACTGGGAAACACTTACCG 588
 b 181 CGAAAGTAAACCTGAAAGTGAAGAGACTTGTCTTTGACAAAGTACACTGGGAAACACTTACCG 240
 Y 589 AGTGGGTGACACTTATGAGCGTCTTAAGAGCTCCATGATCTGGAGCTGTACCTGGCATCGG 648
 b 241 AGTGGGTGACACTTATGAGCGTCTTAAGAGCTCCATGATCTGGAGCTGTACCTGGCATCGG 300
 Y 649 GGCTGGCGGAGGAGAAATAGCTGTACCTATCGCAAAACCGCTGCAATGAAGGGGGTCACTC 708
 b 301 GGCTGGCGGAGGAGAAATAGCTGTACCTATCGCAAAACCGCTGCAATGAAGGGGGTCACTC 360
 Y 709 CTACAGATTGGTGCACACCTGGAGGAGACACATGAGACTGGTGGTTACATGTTAGAGTG 768

Db 361 CTACAGATTGGTGCACCTGGAGGAGACCAATGAGACTGGTGGTTACATGTTAGAGTG 420
 Y 769 TGTGTCTCTTGTAAATGGAAAGAGGAATGGACCTGCAAGCCCATAGCTGAGAAAGTGT 828
 Db 421 TGTGTCTCTTGTAAATGGAAAGAGGAATGGACCTGCAAGCCCATAGCTGAGAAAGTGT 480
 Y 829 TGATCATGCTGTGGGACTTCTATGTGCTGGAGAAACGTGGAGAAACCTTACCAAGG 888
 Db 481 TGATCATGCTGTGGGACTTCTATGTGCTGGAGAAACGTGGAGAAACCTTACCAAGG 540
 Y 889 CTGGATGATGGTAGATTGTACTTGTCTGGAGAAAGCAGCGGAGCATCACTTGACATTC 948
 Db 541 CTGGATGATGGTAGATTGTACTTGTCTGGAGAAAGCAGCGGAGCATCACTTGACATTC 600
 Y 949 TAGAATAGATGCAACGATCAGACACACAGGACATCCTATAGAAATGGAGACACCTGGAG 1008
 Db 601 TAGAATAGATGCAACGATCAGACACACAGGACATCCTATAGAAATGGAGACACCTGGAG 560
 Y 1009 CAAGAAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGCACAGGCAACGGCGGAGAGA 1068
 Db 661 CAAGAAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGCACAGGCAACGGCGGAGAGA 720
 Y 1069 GTGGAAGTGTGAGAGGACACCTCTCTGTCAGACACATCGAGCGGATCTGGCCCTTCAC 1128
 Db 721 GTGGAAGTGTGAGAGGACACCTCTCTGTCAGACACATCGAGCGGATCTGGCCCTTCAC 780
 Y 1129 CGATGTTCTGTGAGCTGTTTACCAACCGCAGCCCTCACCCCGAGCTCTCCCTATGCGCA 1188
 Db 781 CGATGTTCTGTGAGCTGTTTACCAACCGCAGCCCTCACCCCGAGCTCTCCCTATGCGCA 840
 Y 1189 CTGTGTACAGACAGTGGTGTGTCTCTGTGGGATGTCAGTGTCTGAGACACAAGG 1248
 Db 841 CTGTGTACAGACAGTGGTGTGTCTCTGTGGGATGTCAGTGTCTGAGACACAAGG 900
 Y 1249 AATAAGCAAAATGCTTGTGACGTGCTGGCAACGGAGTCAGCTGCCAAGAGACAGCTGT 1308
 Db 901 AATAAGCAAAATGCTTGTGACGTGCTGGCAACGGAGTCAGCTGCCAAGAGACAGCTGT 960
 Y 1309 AACCAGACTTACGGTGGCAACTCAAAATGGAGAGCCATGTGTCTTACCATTACCTACAA 1368
 Db 961 AACCAGACTTACGGTGGCAACTTAAATGGAGAGCCATGTGTCTTACCATTACCTACAA 1020
 Y 1369 TGGCAGGACGTTCTACTCTGACACCAAGAGGCGCAGAGGAGCATCTTTGGTGCAG 1428
 Db 1021 TGGCAGGACGTTCTACTCTGACACCGAGGAGGCGCAGAGCATCTTTGGTGCAG 1080
 Y 1429 CACAACCTCGAAATATGAGCAGGACAGAAATCTCTTTCTGCACAGACACACTGTTTT 1488
 Db 1081 CACAACCTCGAAATATGAGCAGGACAGAAATCTCTTTCTGCACAGACACACTGTTTT 1140
 Y 1489 GGTTTCAGACTCGAGGAGAAATTCGAATGCTGTGCTGCTTCCCTTCCCTATACAA 1548
 Db 1141 GGTTTCAGACTCGAGGAGAAATTCGAATGCTGTGCTGCTTCCCTTCCCTATACAA 1200
 Y 1549 CAACCAAAATATACCTGATTCGACTTCTGAGGCGAAGAGACAACATGAAGTGGTGTGG 1608
 Db 1201 CAACCAAAATATACCTGATTCGACTTCTGAGGCGAAGAGACAACATGAAGTGGTGTGG 1260
 Y 1609 GACCACACAGAACTATGATCCGACAGAAATTTGGTCTGCTGCCCATGCTGCCACCA 1668
 Db 1261 GACCACACAGAACTATGATCCGACAGAAATTTGGTCTGCTGCCCATGCTGCCACCA 1320
 Y 1669 GGAAATCTGCACAAACCAATGAAGGGTCTATGTAACCGCATTTGGAGATCAGTGGGATGA 1728
 Db 1321 GGAAATCTGCACAAACCAATGAAGGGTCTATGTAACCGCATTTGGAGATCAGTGGGATGA 1380
 Y 1729 GCATGACATGGGTACATGATGAGGTGCACTGTGTGGGAATGGTCTGGGGAATGGAC 1788
 Db 1381 GCATGACATGGGTACATGATGAGGTGCACTGTGTGGGAATGGTCTGGGGAATGGAC 1440
 Y 1789 ATGCAATTCCTCTCTGAGCTTTCAGATCAGTGTGTTGATGACATCACTTACATGT 1848

4009 TTACTGTCAAGGATGACAAAGGAAGTGTCCCTATCTCTGATACCATCATCCAGCTGT 4068
3661 TTACTGTCAAGGATGACAAAGGAAGTGTCCCTATCTCTGATACCATCATCCAGCTGT 3720
4069 TCCTCCTCCACTGACCTGCGGATTCACCAACATTTGGTCCAGACACCATGCGTGTCACTG 4128
3721 TCCTCCTCCACTGACCTGCGGATTCACCAACATTTGGTCCAGACACCATGCGTGTCACTG 3780
4129 GGCTCCACCCCCATCCATGTATTTAAACAACTTCTGTGTGGTGTACTACCTGTGAAAAA 4188
3781 GGCTCCACCCCCATCCATGTATTTAAACAACTTCTGTGTGGTGTACTACCTGTGAAAAA 3840
4189 TGAGAGAGATGTTGAGAGTGTCTCAATTTCTCTTCCAGACATGAGTGTCTTAAACAAA 4248
3841 TGAGAGAGATGTTGAGAGTGTCTCAATTTCTCTTCCAGACATGAGTGTCTTAAACAAA 3900
4249 TCTCCTGCTGTTACAGAAATGTTAGTGTGTCTCCAGTGTCTACGAAACAAATGAGAG 4308
3901 TCTCCTGCTGTTACAGAAATGTTAGTGTGTCTCCAGTGTCTACGAAACAAATGAGAG 3960
4309 CACACTCTTAGAGAGACAGAAACAGGCTCTGATTTCCCAACTGCGCATTTGACTTTTC 4368
3961 CACACTCTTAGAGAGACAGAAACAGGCTCTGATTTCCCAACTGCGCATTTGACTTTTC 4020
4369 TGATATTTACTGCAACTCTTTTACTGTGTGCACTGGATTTCTCTCGAGCCCACTCACTGG 4428
4021 TGATATTTACTGCAACTCTTTTACTGTGTGCACTGGATTTCTCTCGAGCCCACTCACTGG 4080
4429 CTACAGGATCCGCAATCATCCCGAGACCTTCACTGGAGAGCTCGAGAGATCCGGTGC 4488
4081 CTACAGGATCCGCAATCATCCCGAGACCTTCACTGGAGAGCTCGAGAGATCCGGTGC 4140
4489 CCACCTCTCGAAATTCATCACCCTCACCACCTCACTCCAGGACAGAGTATGTGTGAG 4548
4141 CCACCTCTCGAAATTCATCACCCTCACCACCTCACTCCAGGACAGAGTATGTGTGAG 4200
4549 CATGCTGTCTTAATGGAGAGAGAGAGTCCCTTATTTGATTTGGCCCAACATCAACAGT 4608
4201 CATGCTGTCTTAATGGAGAGAGAGAGTCCCTTATTTGATTTGGCCCAACATCAACAGT 4260
4609 TTCTGATGTTCCGAGGACCTGGAAAGTGTGCTGCGACCCCTCCAGCTACTGATCAG 4668
4261 TTCTGATGTTCCGAGGACCTGGAAAGTGTGCTGCGACCCCTCCAGCTACTGATCAG 4320
4669 CTGGAGATGCTCTGTGTGTACAGTGTAGATATTTACAGGATCACTTACGAGGAAACAGGAGG 4728
4321 CTGGAGATGCTCTGTGTGTACAGTGTAGATATTTACAGGATCACTTACGAGGAAACAGGAGG 4380
4729 AAATAGCCTGTCCAGGAGTTCACCTGTGCTGGAGCAAGTCTACAGTCACTACCATCAGCGG 4788
4381 AAATAGCCTGTCCAGGAGTTCACCTGTGCTGGAGCAAGTCTACAGTCACTACCATCAGCGG 4440
4789 CTTTAAACCTGGAGTTGATTTATACATCTCTGTGTATGTGTCTCACTGGCGTGGAGACAG 4848
4441 CTTTAAACCTGGAGTTGATTTATACATCTCTGTGTATGTGTCTCACTGGCGTGGAGACAG 4500
4849 CCCCAGACGACGACCAATTTCCATTAATTTACGAAACAGAAATTTGACAAACCATCCCA 4908
4501 CCCCAGACGACGACCAATTTCCATTAATTTACGAAACAGAAATTTGACAAACCATCCCA 4560
4909 GATGCAAGTGAACCGATGTTTACGACAAACAGCAATTTAGTGTCAAGTGGCTGCTTCAAGTTT 4968
4561 GATGCAAGTGAACCGATGTTTACGACAAACAGCAATTTAGTGTCAAGTGGCTGCTTCAAGTTT 4620
4969 CCTGTGTCTGTTTACAGAGTACCACTCTCCCAAAATGGACAGGACCAACAAACAAAC 5028
4621 CCTGTGTCTGTTTACAGAGTACCACTCTCCCAAAATGGACAGGACCAACAAACAAAC 4680
5029 TAAACTGCAGGTCCTAGATCAACAGAAATGACTATTTGAGGCTTTCAGGCCCAAGTGA 5088
4681 TAAACTGCAGGTCCTAGATCAACAGAAATGACTATTTGAGGCTTTCAGGCCCAAGTGA 4740
5089 GTATGTGGTGTAGTGTCTATGCTCAGAAATCCAGCGGAGAGAGTCAAGCTCTGGTTCAGAC 5148

4741 GTATGTGGTGTAGTGTCTATGCTCAGAAATCCAGCGGAGAGAGTCAAGCTCTGGTTCAGAC 4800
5149 TGCAAGTAAACCAATGATGTGCGCTTAAAGGACTGGCATTTCACTGATGTGGATGTGCAATTC 5208
4801 TGCAAGTAAACCAATGATGTGCGCTTAAAGGACTGGCATTTCACTGATGTGGATGTGCAATTC 4860
5209 CATCAAAATTTGCTTGGAAAGCCCAAGGGGCAAGTTTCCAGGTACAGGGTACACCTACTC 5268
4861 CATCAAAATTTGCTTGGAAAGCCCAAGGGGCAAGTTTCCAGGTACAGGGTACACCTACTC 4920
5269 GAGCCTGAGGATGGAATCCATGAGCTATTCCTCTGCACTGTATGTGTGAAGAACACTGTC 5328
4921 GAGCCTGAGGATGGAATCCATGAGCTATTCCTCTGCACTGTATGTGTGAAGAACACTGTC 4980
5329 AGAGCTGCAAGGCTCAGACCGGTTTCTGAGTACACAGTCACTGTGTGTGCTTGCACGA 5388
4981 AGAGCTGCAAGGCTCAGACCGGTTTCTGAGTACACAGTCACTGTGTGTGCTTGCACGA 5040
5389 TGATATGGAGAGCCAGCCCTGATTGGAAACCCAGTCCACAGCTATTTCTGTGCAACCACTGA 5448
5041 TGATATGGAGAGCCAGCCCTGATTGGAAACCCAGTCCACAGCTATTTCTGTGCAACCACTGA 5100
5449 CTTGAAGTTTCACTCAGGTCAACCCACAGCTGAGCGCCAGTGGACACCACTCACTGTC 5508
5101 CTTGAAGTTTCACTCAGGTCAACCCACAGCTGAGCGCCAGTGGACACCACTCACTGTC 5160
5509 TCAGCTCACTGGATATCGAGTGGGTGACCCCAAGGAGAGACCGGACCACTCACTGTC 5568
5161 TCAGCTCACTGGATATCGAGTGGGTGACCCCAAGGAGAGACCGGACCACTCACTGTC 5220
5569 AATCAACCTTGTCTGAGAGCTCATCCGTTGTTGATCAGGACTATGTTGGCCACCA 5628
5221 AATCAACCTTGTCTGAGAGCTCATCCGTTGTTGATCAGGACTATGTTGGCCACCA 5280
5629 ATATGAAGTCAAGTGTCTATGCTCTTAAGGACACTTTTGAACGACAGCAGCTCAGGAGT 5688
5281 ATATGAAGTCAAGTGTCTATGCTCTTAAGGACACTTTTGAACGACAGCAGCTCAGGAGT 5340
5689 TGTCAACCTCTGGAGATGTGAGCCCAACAGAGGGCTGTTGACAGATGCTACTGA 5748
5341 TGTCAACCTCTGGAGATGTGAGCCCAACAGAGGGCTGTTGACAGATGCTACTGA 5400
5749 GACCACCATCACCATTAGCTGGAGAACCAAGACTGAGAGCACTCACTGGCTCCAACTTGA 5808
5401 GACCACCATCACCATTAGCTGGAGAACCAAGACTGAGAGCACTCACTGGCTCCAACTTGA 5460
5809 TGCGTTTCCAGCAATGCGCAGACTCCAAATCCAGAGAACCATCAAGCCAGATGTCAAG 5868
5461 TGCGTTTCCAGCAATGCGCAGACTCCAAATCCAGAGAACCATCAAGCCAGATGTCAAG 5520
5869 CTACACCATCAGCTGGCTTACAAACGAGCACTGACTCAAGATCTACTGTACACCTTGA 5928
5521 CTACACCATCAGCTGGCTTACAAACGAGCACTGACTCAAGATCTACTGTACACCTTGA 5580
5929 TGCAATGTCTCGAGTCTCCCTGTGTCTGATGCAAGCTTCCATGTCATTTGATGCACTC 5988
5581 TGCAATGTCTCGAGTCTCCCTGTGTCTGATGCAAGCTTCCATGTCATTTGATGCACTC 5640
5989 CAACCTGCGTTTCTGCGCACCAACCCCAATTCCTTGTGTATCATGCGAGCGCCACG 6048
5641 CAACCTGCGTTTCTGCGCACCAACCCCAATTCCTTGTGTATCATGCGAGCGCCACG 5700
6049 TGCCAGGATTAACCGGCTTACATCATCAAGTATGAGAGCTGGGTCTCTCCCGAGAGAGT 6108
5701 TGCCAGGATTAACCGGCTTACATCATCAAGTATGAGAGCTGGGTCTCTCCCGAGAGAGT 5760
6109 GGTCCCTCGCGCCCGCTGCTGTCTCAGAGGCTACTATTTACTGGCTGTGAACCGGAG 6168
5761 GGTCCCTCGCGCCCGCTGCTGTCTCAGAGGCTACTATTTACTGGCTGTGAACCGGAG 5820
6169 CGAATATACAAATTTATGCTCATTTGCTGAAAGATAATCAGAGAGCGAGCCCTGATTGG 6228

5821 CGAATATACAAATTTATGTCATTTGCCCTGGAAGATAATATCAGAGAGCGAGCCCTGATGG 5880
6229 AAGGAAAAAGACAGACGAGCTTCCCAACTGGTAACCCCTTCCACACCCCAATCTTTCATGG 6288
5881 AAGGAAAAAGACAGACGAGCTTCCCAACTGGTAACCCCTTCCACACCCCAATCTTTCATGG 5940
6289 ACCAGAGATCTTGATGTTCTTCCACAGTTCAAAGACCCCTTTCGTCACCAACCTGG 6348
5941 ACCAGAGATCTTGATGTTCTTCCACAGTTCAAAGACCCCTTTCGTCACCAACCTGG 6000
6349 GTATGACACTGGAATGGTATTCAGCTTCTCTGGCACTTCTGGTCAGCAACCCCAAGTGTGG 6408
6001 GTATGACACTGGAATGGTATTCAGCTTCTCTGGCACTTCTGGTCAGCAACCCCAAGTGTGG 6060
6409 GCAACAAATGATCTTTGAGGAACATGTTTATGAGCGGACACACCCGCCCAACAGGCCAC 6468
6061 GCAACAAATGATCTTTGAGGAACATGTTTATGAGCGGACACACCCGCCCAACAGGCCAC 6120
6469 CCCATAGGCGATAGGCCAAGACCATACCCGCCGAATGTAGGCAAGAACTCTCTCTCA 6528
6121 CCCATAGGCGATAGGCCAAGACCATACCCGCCGAATGTAGGCAAGAACTCTCTCTCA 6180
6529 GACAACTCTCATGGGCCCCCATCTCCAGGACATCTCTGAGTACATCATCTATGTCATCC 6588
6181 GACAACTCTCATGGGCCCCCATCTCCAGGACATCTCTGAGTACATCATCTATGTCATCC 6240
6589 TGTGGCACTGATGAAGAACCTTTACAGTTCAGGTTCTCGAACTTCTACCACTGCGCAC 6648
6241 TGTGGCACTGATGAAGAACCTTTACAGTTCAGGTTCTCGAACTTCTACCACTGCGCAC 6300
6649 TCTGACAGGCTCTACAGAGTGCCACTTACAACTATAGTGAGGAGCTGGAAGACCA 6708
6301 TCTGACAGGCTCTACAGAGTGCCACTTACAACTATAGTGAGGAGCTGGAAGACCA 6360
6709 GCAGAGCATTAAGTTCCGGAGAGGTTGTTACCGTGGGCACTCTGTCAACAGAGCTT 6768
6361 GCAGAGCATTAAGTTCCGGAGAGGTTGTTACCGTGGGCACTCTGTCAACAGAGCTT 6420
6769 GAACCAACTACGAGTACTCGTCTTTGACCCCTTACACAGTGTCCCATATGCGGTGG 6828
6421 GAACCAACTACGAGTACTCGTCTTTGACCCCTTACACAGTGTCCCATATGCGGTGG 6480
6829 AGATGAGTGGAAAGAACTGATCAGAGCTTTAACTGTTGTCAGTGTGCTTACGCTT 6888
6481 AGATGAGTGGAAAGAACTGATCAGAGCTTTAACTGTTGTCAGTGTGCTTACGCTT 6540
6889 TGGAGTGGTCACTTTTACAGTGTGATCTATCTAGATGTTGCCATGACAAATGGTGTGAAC 6948
6541 TGGAGTGGTCACTTTTACAGTGTGATCTATCTAGATGTTGCCATGACAAATGGTGTGAAC 6600
6949 CAAGTTGAGAGAGTGGGACCGTCCAGGAGAAATGGCCAGATGATGAGTGCACATG 7008
6601 CAAGTTGAGAGAGTGGGACCGTCCAGGAGAAATGGCCAGATGATGAGTGCACATG 6660
7009 TCTTGGGACGGAAGAGAAATTCAGTGTGACCCCTCATGAGCAACGTTGTACGATGA 7068
6661 TCTTGGGACGGAAGAGAAATTCAGTGTGACCCCTCATGAGCAACGTTGTACGATGA 6720
7069 TGGGAAGCATACCACTAGGAGACAGTGGAGAAATCTCGGTGCATTTGCTC 7128
6721 TGGGAAGCATACCACTAGGAGACAGTGGAGAAATCTCGGTGCATTTGCTC 6780
7129 CTGCACATGCTTTGAGGCGCAGCGGGTGGCGCTGTGACAACTGCCGACACCTGGGGG 7188
6781 CTGCACATGCTTTGAGGCGCAGCGGGTGGCGCTGTGACAACTGCCGACACCTGGGGG 6840
7189 TGAACCCAGTCCGGAAGGCACTACTGGCAGTCTTACAAACAGTATCTTCAGATACCA 7248
6841 TGAACCCAGTCCGGAAGGCACTACTGGCAGTCTTACAAACAGTATCTTCAGATACCA 6900
7249 TCAGAGAACAAACATAATGTTAAATGCGCAATGAGTGTCTTCATGCCCTTTAGATGTACA 7308
6901 TCAGAGAACAAACATAATGTTAAATGCGCAATGAGTGTCTTCATGCCCTTTAGATGTACA 6960

7309 GGCTGACAGAGAAGATTCCTCCAGAGTAATAATCATCTTTCCAAATCCAGAGGAACAAGCATGT 7368
6961 GGCTGACAGAGAAGATTCCTCCAGAGTAATAATCATCTTTCCAAATCCAGAGGAACAAGCATGT 7020
7369 CTCTCTGCCAAGATCCATCTAAACTGGAGTGTATGTTAGCAGACCCAGCTTAGAGTTCTTC 7428
7021 CTCTCTGCCAAGATCCATCTAAACTGGAGTGTATGTTAGCAGACCCAGCTTAGAGTTCTTC 7080
7429 TTTCTTTCTTAAAGCCCTTGTCTCTGGAGGAAGTCTCCAGCTTCAGCTCAACTCACAGCT 7488
7081 TTTCTTTCTTAAAGCCCTTGTCTCTGGAGGAAGTCTCCAGCTTCAGCTCAACTCACAGCT 7140
7489 TCTCAAGCATCACCTCGGAGTTTCTGAGGGTTTCTCATAAATGAGGCTGCACACT 7548
7141 TCTCAAGCATCACCTCGGAGTTTCTGAGGGTTTCTCATAAATGAGGCTGCACACT 7200
7549 GGCTGTTCTGCTCGAAGTATTCAATACCGCTCAGTATTTTAAATGAAAGTGAATCTTAAGA 7608
7201 GGCTGTTCTGCTCGAAGTATTCAATACCGCTCAGTATTTTAAATGAAAGTGAATCTTAAGA 7260
7609 TTTGTTTGGGATCAATAGGAAGCATATGAGCCCAACCAAGATGCAAAATGTTTGAAT 7668
7261 TTTGTTTGGGATCAATAGGAAGCATATGAGCCCAACCAAGATGCAAAATGTTTGAAT 7320
7669 GATATGACCAAAATTTTAAAGTAGGAAGTCAACCAACCACTTCTGCTTCACTTAAAGTGT 7728
7321 GATATGACCAAAATTTTAAAGTAGGAAGTCAACCAACCACTTCTGCTTCACTTAAAGTGT 7380
7729 CTGGCCCGCAACTGTTAGGAACAGCATGATCTTGTACTGTGATATTTTAAATATCCA 7788
7381 CTGGCCCGCAACTGTTAGGAACAGCATGATCTTGTACTGTGATATTTTAAATATCCA 7440
7789 CAGTACT 7795
7441 CAGTACT 7447

RESULT 10

US-10-236-031B-69
; Sequence 69, Application US/10236031B
; Publication No. US20030219760A1
; GENERAL INFORMATION:
; APPLICANT: Gordon, Gavin J.
; APPLICANT: Jensen, Roderick V.
; APPLICANT: Gullens, Steven R.
; TITLE OF INVENTION: Diagnostic and Prognostic Tests
; FILE REFERENCE: B0801/70265 (JRV/JRV)
; CURRENT APPLICATION NUMBER: US/10/236,031B
; CURRENT FILING DATE: 2002-09-05
; PRIOR APPLICATION NUMBER: US 60/317,389
; PRIOR FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: US 60/407,431
; PRIOR FILING DATE: 2002-08-30
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 69
; LENGTH: 7680
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-236-031B-69

Query Match 95.2%; Score 7423; DB 15; Length 7680;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 7432; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 349 GAAGACAAGAGGAGGCTCAGCAATGTTTACAGCCCAAGTCCCGGTGGCTGTCAGTCA 408
Db 1 GAAGACAAGAGGAGGCTCAGCAATGTTTACAGCCCAAGTCCCGGTGGCTGTCAGTCA 60
Qy 409 AAGCAAGCCGGTGTGTTATGACAAATGGAATACTATCAGATAAATCAACAGTGGAGCG 468

2629	Y	TGTAATGTCTATCAGATATCTGAGATGGGAGCAGAGTTTGATCTCTGTCTTACTTCA	2688
2281	b	TGTAATGTCTATCAGATATCTGAGATGGGAGCAGAGTTTGATCTCTGTCTTACTTCA	2340
2689	Y	AACACAGCGCTGATGCCCCCTCTGACCCGACTGTGGACCAAGTTGATGACACCTCAAT	2748
2341	b	AACACAGCGCTGATGCCCCCTCTGACCCGACTGTGGACCAAGTTGATGACACCTCAAT	2400
2749	Y	TGTTCTTGGCTGGAGCAGACCCGAGCTCCCATCAGGGTACAGATAGTCTATTGCCC	2808
2401	b	TGTTCTTGGCTGGAGCAGACCCGAGCTCCCATCAGGGTACAGATAGTCTATTGCCC	2460
2809	Y	ATCAGTAGAAGTAGCAGACAGAACTCAACCTTCTGAAACTGCAAACTCCGTCACTT	2868
2461	b	ATCAGTAGAAGTAGCAGACAGAACTCAACCTTCTGAAACTGCAAACTCCGTCACTT	2520
2869	Y	CAGTGACATTGCAACCTGGTGTTCAGTATAACATCACTATCTATGTTGGAAAGAAATCA	2928
2521	b	CAGTGACATTGCAACCTGGTGTTCAGTATAACATCACTATCTATGTTGGAAAGAAATCA	2580
2929	Y	AGAAAGTAGACCTGTTTGTCAATTCAACAGAAACCACTGGCACTCCACGCTCAGATACAGT	2988
2581	b	AGAAAGTAGACCTGTTTGTCAATTCAACAGAAACCACTGGCACTCCACGCTCAGATACAGT	2640
2989	Y	GCCTCTCCACGGACCTGCAGTTTGTGAAAGTGACAGACGTGAAGGTCAACATCATGTG	3048
2641	b	GCCTCTCCACGGACCTGCAGTTTGTGAAAGTGACAGACGTGAAGGTCAACATCATGTG	2700
3049	Y	GACACCGCTGAGATGACGCGGTACCGGTGAGATGTATCCCCGCTCAACTGCTGC	3108
2701	b	GACACCGCTGAGATGACGCGGTACCGGTGAGATGTATCCCCGCTCAACTGCTGC	2760
3109	Y	TGGCAGACGGGCAGAGCTGCCCATCAGCAGGAACACTTTGCAGAAAGTCAACGGGGT	3168
2761	b	TGGCAGACGGGCAGAGCTGCCCATCAGCAGGAACACTTTGCAGAAAGTCAACGGGGT	2820
3169	Y	GTCCCTTGGGTCACTATTACTTCAAAGCTTTTGAGTGAAGCCATGGGAGGAGAGCAA	3228
2821	b	GTCCCTTGGGTCACTATTACTTCAAAGCTTTTGAGTGAAGCCATGGGAGGAGAGCAA	2880
3229	Y	GCCTCTGACTGCTCAACAGACACCAACTGGATGCTCCCACTAACTCCAGTTTCTCAA	3288
2881	b	GCCTCTGACTGCTCAACAGACACCAACTGGATGCTCCCACTAACTCCAGTTTCTCAA	2940
3289	Y	TGAACTGATTCTACTGTCTGTGTAGATGGATCCACTTCGGGCCAGATACAGGATA	3348
2941	b	TGAACTGATTCTACTGTCTGTGTAGATGGATCCACTTCGGGCCAGATACAGGATA	3000
3349	Y	CCGACTGAACGTGGGCTTTACCCGAGAGAGCAGCCAGGCTCAAGTTGGGTCCCTC	3408
3001	b	CCGACTGAACGTGGGCTTTACCCGAGAGAGCAGCCAGGCTCAAGTTGGGTCCCTC	3060
3409	Y	TGTCTCCAAGTACCCACTGAGGAATCTGAGCCTGTCATCTGAGTACACCGTATCCCTCGT	3468
3061	b	TGTCTCCAAGTACCCCTTGAAGGAATCTGAGCCTGTCATCTGAGTACACCGTATCCCTCGT	3120
3469	Y	GGCATAAAGGGCAACAAGAGAGCCCAAGCCACTGGAGTCTTTACCACTGAGGCC	3528
3121	b	GGCATAAAGGGCAACAAGAGAGCCCAAGCCACTGGAGTCTTTACCACTGAGGCC	3180
3529	Y	TGGAGCTCTATTTCACCTTTACAACACCGAGGTGACTGAGACCAACCATTTGTATCAATG	3588
3181	b	TGGAGCTCTATTTCACCTTTACAACACCGAGGTGACTGAGACCAACCATCTGTATCAATG	3240
3589	Y	GACGCTGTCTCAAGAAATGGTTTTTAAGCTGGGTGTAAGACCAAGCCAGGGAGGAGGC	3648
3241	b	GACGCTGTCTCAAGAAATGGTTTTTAAGCTGGGTGTAAGACCAAGCCAGGGAGGAGGC	3300
3649	Y	ACCACAGAGTGACTTCAGACTCAGGAACATCGTTGTGTCGGCTTGACTCCAGAGT	3708
3301	b	ACCACAGAGTGACTTCAGACTCAGGAACATCGTTGTGTCGGCTTGACTCCAGAGT	3360
3709	Y	AGAATAGCTACACCATCCAAGTCTCTGAGATGGCAGGAAAGAGATCGCCAAATTTG	3768

Db	3361	AGAAATACGCTTACCACTCAAGTCTCTGAGATGGACAGGAAGAGATGGCCATTGT	3422
Qy	3769	AAACAAAGTGGTGACACCAATTGTTCTCCACCAAACTTTGCATCTGGAGCAAAACCTCGA	3828
Db	3421	AAACAAAGTGGTGACACCAATTGTTCTCCACCAAACTTTGCATCTGGAGCAAAACCTCGA	3480
Qy	3829	CATCTGAGTGCTCACAGTCTCCTGGAGAGGAGCACCACCCAGACATTACTGTTATTAG	3888
Db	3481	CATCTGAGTGCTCACAGTCTCCTGGAGAGGAGCACCACCCAGACATTACTGTTATTAG	3540
Qy	3889	AAATTACCAAAACCCCTACAAAACCGCCAGCAGGGAAATTTCTTTGGAAAGATGTCCTCATGC	3948
Db	3541	AAATTACCAAAACCCCTACAAAACCGCCAGCAGGGAAATTTCTTTGGAAAGATGTCCTCATGC	3600
Qy	3949	TGATCAGAGCTCCTGCACATTTTGATTAACCTGAGTCCGGCCTGGAGTACAAATGTCAGTGT	4008
Db	3601	TGATCAGAGCTCCTGCACATTTTGATTAACCTGAGTCCGGCCTGGAGTACAAATGTCAGTGT	3660
Qy	4009	TTACACTGTCAAGGATGACAAGGAAAGTGTCCCTATCTCTGATACCATCATCCACAGCTGT	4068
Db	3661	TTACACTGTCAAGGATGACAAGGAAAGTGTCCCTATCTCTGATACCATCATCCACAGCTGT	3720
Qy	4069	TCCTCTCCCACTGACCTGGATTCACCAACATTTGGTTCAGACACCATGCGTGCACCTG	4128
Db	3721	TCCTCTCCCACTGACCTGGATTCACCAACATTTGGTTCAGACACCATGCGTGCACCTG	3780
Qy	4129	GGCTCCACCCCACTCAATTGATTAACCAACTTCTGGTGCCTTACTCACTGTGAAAAA	4188
Db	3781	GGCTCCACCCCACTCAATTGATTAACCAACTTCTGGTGCCTTACTCACTGTGAAAAA	3840
Qy	4189	TGAGGAAGATGTTTCAGAGTGTCAATTTCTCCTTCAGACAATGCAGTGTCTTTAACAAA	4248
Db	3841	TGAGGAAGATGTTTCAGAGTGTCAATTTCTCCTTCAGACAATGCAGTGTCTTTAACAAA	3900
Qy	4249	TCCTCTGCCTGGTACAGAAATATGTAGTGAGTGCTCCAGTGCTACGAACAAATGAGAG	4308
Db	3901	TCCTCTGCCTGGTACAGAAATATGTAGTGAGTGCTCCAGTGCTACGAACAAATGAGAG	3960
Qy	4309	CACACCTCTTAGAGGAAGACAGAAAAAGGTCTTGATTTCCCAACTGGCAATGCACTTTTC	4368
Db	3961	CACACCTCTTAGAGGAAGACAGAAAAAGGTCTTGATTTCCCAACTGGCAATGCACTTTTC	4020
Qy	4369	TGATATTACTGCCAACTCTTTTACTGTGCACTGGATTGCTCCTCGAGCACCACTACTGG	4428
Db	4021	TGATATTACTGCCAACTCTTTTACTGTGCACTGGATTGCTCCTCGAGCCACCATCTGG	4080
Qy	4429	CTACAGATCGCCATCATCCGAGCACTTCAGTGGAGACCTCGAAGAATCGSGTCC	4488
Db	4081	CTACAGATCGCCATCATCCGAGCACTTCAGTGGAGACCTCGAAGAATCGSGTCC	4140
Qy	4489	CCACTCTCGGAATTCATACCTCTCACCAACCTCACTCCAGGCACAGATATGTGTCAG	4548
Db	4141	CCACTCTCGGAATTCATACCTCTCACCAACCTCACTCCAGGCACAGATATGTGTCAG	4200
Qy	4549	CATGTTGCTCTTAATGGCAGAGGAAGTCCCTTATTTGATTGGCCAAACAATCAACAGT	4608
Db	4201	CATGTTGCTCTTAATGGCAGAGGAAGTCCCTTATTTGATTGGCCAAACAATCAACAGT	4260
Qy	4609	TTCTGATGTTCCGAGGACCTGGAAGTTGTTGTCGACCCCAACAGCCTACTGTATCAG	4668
Db	4261	TTCTGATGTTCCGAGGACCTGGAAGTTGTTGTCGACCCCAACAGCCTACTGTATCAG	4320
Qy	4669	CTGGATGCTCTCTGCTGCAGTGAGATATTACAGGATCACTTACGGAGAAAACAGAGG	4728
Db	4321	CTGGATGCTCTCTGCTGCAGTGAGATATTACAGGATCACTTACGGAGAAAACAGAGG	4380
Qy	4729	AAATAGCCCTGTCCAGGAGTTCACTGTCCCTGGAGCAAGTCTACAGCTACCATCAGCGG	4788
Db	4381	AAATAGCCCTGTCCAGGAGTTCACTGTCCCTGGAGCAAGTCTACAGCTACCATCAGCGG	4440
Qy	4789	CCTTAAACCTGGAGTTGATTATACCATCACTGTGTATGCTGTCACTGGCCGTGGAGACAG	4848

b 4441 CCTTAAACCTGGAGTGTGATTATACCATCACTGTGTATGCTCTCACTGGCGTGGAGACAG 4500
 y 4849 CCCCAGACAGCAAGCCAAATTCATTAATTAACGAAACAGAAATTAACAAACCATCCCA 4908
 b 4501 CCCCAGACAGCAAGCCAAATTCATTAATTAACGAAACAGAAATTAACAAACCATCCCA 4560
 y 4909 GATCAAGTGCAGGATGTTCCAGCAACAGCAATAGTGTCAAGTGGCTGCTTCAAGTTC 4968
 b 4561 GATCAAGTGCAGGATGTTCCAGCAACAGCAATAGTGTCAAGTGGCTGCTTCAAGTTC 4620
 y 4969 CCGCTGTACTGTGTACAGAGTAACCAACCACTCCCAAAATGGAACAGAGCAACAAAC 5028
 b 4621 CCGCTGTACTGTGTACAGAGTAACCAACCACTCCCAAAATGGAACAGAGCAACAAAC 4680
 y 5029 TAAACTGCAGGTCAGATCAACAGCAATAGTGTCAAGGCTTGCAGCCACAGTGA 5088
 b 4681 TAAACTGCAGGTCAGATCAACAGCAATAGTGTCAAGGCTTGCAGCCACAGTGA 4740
 y 5089 GTATGTGGTGTAGTGTCTATGCTCAGAATCCAAAGCGAGAGTCAAGCTCTGGTTCAGAC 5148
 b 4741 GTATGTGGTGTAGTGTCTATGCTCAGAATCCAAAGCGAGAGTCAAGCTCTGGTTCAGAC 4800
 y 5149 TGCAGTACCAACATGATCGCCCTAAGGACTGGCATTCACTGATGTGGATGTCGATTC 5208
 b 4801 TGCAGTACCAACATGATCGCCCTAAGGACTGGCATTCACTGATGTGGATGTCGATTC 4860
 y 5209 CATCAAAATGTCTTGGAAAGCCCAAGGGCAAGTTTCCAGGTACAGGGTGAACCTACTC 5268
 b 4861 CATCAAAATGTCTTGGAAAGCCCAAGGGCAAGTTTCCAGGTACAGGGTGAACCTACTC 4920
 y 5269 GAGCCCTGAGATGGAATCCATGAGTATTCCTGCACTGATGTGAAGACACTGC 5328
 b 4921 GAGCCCTGAGATGGAATCCATGAGTATTCCTGCACTGATGTGAAGACACTGC 4980
 y 5329 AGAGCTGCAAGGCTCAGACCCGGTTCAGTACACAGTCAAGTGTGGTTCCTTCCACGA 5388
 b 4981 AGAGCTGCAAGGCTCAGACCCGGTTCAGTACACAGTCAAGTGTGGTTCCTTCCACGA 5040
 y 5389 TGATATGGAGAGCCAGCCCTGATGGAAACAGTCCAGCTATTCCTGACCAACTGA 5448
 b 5041 TGATATGGAGAGCCAGCCCTGATGGAAACAGTCCAGCTATTCCTGACCAACTGA 5100
 y 5449 CCTGAAGTTCACCTCAGTCAACCCAGGCTGAGCGCCAGTGGACACCAACCAATGT 5508
 b 5101 CCTGAAGTTCACCTCAGTCAACCCAGGCTGAGCGCCAGTGGACACCAACCAATGT 5160
 y 5509 TCAGCTCACTGGATATCGAGTGGGGTGACCCCAAGGAGAGACCGGACCAATGAAGA 5568
 b 5161 TCAGCTCACTGGATATCGAGTGGGGTGACCCCAAGGAGAGACCGGACCAATGAAGA 5220
 y 5569 AATCAACCTTGCTCCTGACAGCTCATCCGTTGTTGATCAGGACTTATGGTGGCCACCAA 5628
 b 5221 AATCAACCTTGCTCCTGACAGCTCATCCGTTGTTGATCAGGACTTATGGTGGCCACCAA 5280
 y 5629 ATATGAAGTGAAGTGTCTATGCTCTTAAGACACTTTGAACAGCAGACCAAGCTCAGGGAGT 5688
 b 5281 ATATGAAGTGAAGTGTCTATGCTCTTAAGACACTTTGAACAGCAGACCAAGCTCAGGGTGT 5340
 y 5689 TGTCAACACTCTGGAGATGTGAGCCCAACAGAGGGCTCCGTGTGACAGATGCTACTGA 5748
 b 5341 TGTCAACACTCTGGAGATGTGAGCCCAACAGAGGGCTCCGTGTGACAGATGCTACTGA 5400
 y 5749 GACCAACCATCACCATTAGCTGGAGAACCAAGACTGAGACGATCACTGGCTCCAAAGTTGA 5808
 b 5401 GACCAACCATCACCATTAGCTGGAGAACCAAGACTGAGACGATCACTGGCTCCAAAGTTGA 5460
 y 5809 TGGGTTCCAGCAATGGCCAGACTCCATCCAGAGAACATCAAGCCAGATGTCAGAG 5868
 b 5461 TGGCGTTCCAGCAATGGCCAGACTCCATCCAGAGAACATCAAGCCAGATGTCAGAG 5520
 y 5869 CTACACCATCACTGGCTTACAAACAGGCACTGACTACAGATCTACTGTACACCTTGAA 5928
 b 5521 CTACACCATCACTGGCTTACAAACAGGCACTGACTACAGATCTACTGTACACCTTGAA 5580

Qy 5929 TGACATATGCTCGAGTCCCTGTGTGATCATCGAGCTCCCACTGCTGATGACCAATC 5988
 Db 5581 TGACATATGCTCGAGTCCCTGTGTGATCATCGAGCTCCCACTGCTGATGACCAATC 5640
 Qy 5989 CAACTGCGTTCCTGGCCACACACCAATTCCTTGTGTGATCATGCGACCGCCACG 6048
 Db 5641 CAACTGCGTTCCTGGCCACACACCAATTCCTTGTGTGATCATGCGACCGCCACG 5700
 Qy 6049 TGCACAGATTAACCGGTACATCATCAAGTATCAGAGAGCTGCTCTCTCCAGAGAGT 6108
 Db 5701 TGCACAGATTAACCGGTACATCATCAAGTATCAGAGAGCTGCTCTCTCCAGAGAGT 5760
 Qy 6109 GGTCCCTCGCCCGCCCTGGTGTCAAGAGCTACTATTAATGCGCTGGAAACCGGAAAC 6168
 Db 5761 GGTCCCTCGCCCGCCCTGGTGTCAAGAGCTACTATTAATGCGCTGGAAACCGGAAAC 5820
 Qy 6169 CGAATATCAATTTATGTCATTCCTGCTGAAGAAATATCAGAAGAGCGAGCCCTGATTCG 6228
 Db 5821 CGAATATCAATTTATGTCATTCCTGCTGAAGAAATATCAGAAGAGCGAGCCCTGATTCG 5880
 Qy 6229 AAGGAAAAAGACAGACGAGCTTCCCAACTGTGTAACTTCCACACCCCAATCTTCATCG 6288
 Db 5881 AAGGAAAAAGACAGACGAGCTTCCCAACTGTGTAACTTCCACACCCCAATCTTCATCG 5940
 Qy 6289 ACAGAGATCTTGGATGTTCTTCCACAGCTTCAAGAGACCCCTTTCGTCACCCACCTCG 6348
 Db 5941 ACCAGAGATCTTGGATGTTCTTCCACAGCTTCAAGAGACCCCTTTCGTCACCCACCTCG 6000
 Qy 6349 GTATGACACTGGAATATGTTATTCAGCTTCTGTCAGTCTTGTGTCAGCAACCCAGTGTTCG 6408
 Db 6001 GTATGACACTGGAATATGTTATTCAGCTTCTGTCAGTCTTGTGTCAGCAACCCAGTGTTCG 6060
 Qy 6409 GCAACAAATGATCTTTGAGGAACATGTTTAGGCGGACACACCGCCACCAACGGCCAC 6468
 Db 6061 GCAACAAATGATCTTTGAGGAACATGTTTAGGCGGACACACCGCCACCAACGGCCAC 6120
 Qy 6469 CCCCATAAGGCATAGGCCAAGACCACTACCGCCGAAATGATAGGACAAAGAGCTCTCTCTCA 6528
 Db 6121 CCCCATAAGGCATAGGCCAAGACCACTACCGCCGAAATGATAGGACAAAGAGCTCTCTCTCA 6180
 Qy 6529 GAAACCATCTATGGGCCCCCATTCAGGACACTTCTGAGTACATCATTTTCATGTCATCC 6588
 Db 6181 GAAACCATCTATGGGCCCCCATTCAGGACACTTCTGAGTACATCATTTTCATGTCATCC 6240
 Qy 6589 TGTGGCACTGATGAAGAACCCCTTACAGTTCAGGGTTCCTGGAACCTTCTACAGTCCAC 6648
 Db 6241 TGTGGCACTGATGAAGAACCCCTTACAGTTCAGGGTTCCTGGAACCTTCTACAGTCCAC 6300
 Qy 6649 TCTGACAGGCTCAACAGAGTGCACCTACAAACATCATAGTGGAGGCACTGAAAGACCA 6708
 Db 6301 TCTGACAGGCTCAACAGAGTGCACCTACAAACATCATAGTGGAGGCACTGAAAGACCA 6360
 Qy 6709 GCAGAGCATTAAGTTCGGGAAGAGGTTGTACCGTGGCACTCTGTCAACGAAAGCTT 6768
 Db 6361 GCAGAGCATTAAGTTCGGGAAGAGGTTGTACCGTGGCACTCTGTCAACGAAAGCTT 6420
 Qy 6769 GAACCAACCTACGATGACTCGTGTGTCCTTGAACCCCTACAGTGTCCCATTAATGCGCTTCG 6828
 Db 6421 GAACCAACCTACGATGACTCGTGTGTCCTTGAACCCCTACAGTGTCCCATTAATGCGCTTCG 6480
 Qy 6829 AGATGAGTGGGAAGAGTGTGAAATCAGGCTTTAACTGCTGTCAGTGTAGGCTT 6888
 Db 6481 AGATGAGTGGGAAGAGTGTGAAATCAGGCTTTAACTGCTGTCAGTGTAGGCTT 6540
 Qy 6889 TGAAGTGGTCAATTCAGATGTGATTCATCTAGATGGTGCATGACAAATGGTGTGAACCTA 6948
 Db 6541 TGAAGTGGTCAATTCAGATGTGATTCATCTAGATGGTGCATGACAAATGGTGTGAACCTA 6600
 Qy 6949 CAAGATTGGAGAGAGTGGGACCGCTCAGGGAGAAATGGCCAGATGATGAGTGCACATG 7008
 Db 6601 CAAGATTGGAGAGAGTGGGACCGCTCAGGGAGAAATGGCCAGATGATGAGTGCACATG 6660

7009 TCTTGGAAACGGAAGAGAAATCAAGTGTGACCTCATGAGGCAACGTTGTACGATGA 7068
7069 TCTTGGAAACGGAAGAGAAATCAAGTGTGACCTCATGAGGCAACGTTGTACGATGA 6720
7069 TGGGAAGACATACCACTAGGAGAACTGGGAGAGAAATATCTCGGTGCCATTTGCTC 7128
6721 TGGGAAGACATACCACTAGGAGAACTGGGAGAGAAATATCTCGGTGCCATTTGCTC 6780
7129 CTGCACATGCTTTGGAGGCGAGCGGGCTGGCGTGTGACAACTGGCGGAGACCTGGGG 7188
6781 CTGCACATGCTTTGGAGGCGAGCGGGCTGGCGTGTGACAACTGGCGGAGACCTGGGG 6840
7189 TGAACCCAGTCCCGAGGCACTCTGCGGAGTCTCAACAGGATTTCTCAGAGATACCA 7248
6841 TGAACCCAGTCCCGAGGCACTCTGCGGAGTCTCAACAGGATTTCTCAGAGATACCA 6900
7249 TCAGAGAAACAACTAATGTTAATGCCAAATGAGTGTCTTATGCGTTTATGATGATGA 7308
6901 TCAGAGAAACAACTAATGTTAATGCCAAATGAGTGTCTTATGCGTTTATGATGATGA 6960
7309 GGCTGACAGAGATCCCGAGGATTAATCTTTCCCAATCCAGAGGAACAGCATGT 7368
6961 GGCTGACAGAGATCCCGAGGATTAATCTTTCCCAATCCAGAGGAACAGCATGT 7020
7369 CTCTCTGCCAAGATCCATCTAAACTGGAGTGTGTTAGCAGACCCAGCTTTAGAGTTCTTC 7428
7021 CTCTCTGCCAAGATCCATCTAAACTGGAGTGTGTTAGCAGACCCAGCTTTAGAGTTCTTC 7080
7429 TTTCTTTCTTAAGCCCTTGTCTCGGAGGAGTTCTCCAGCTTCAGCTCACTCAGACT 7488
7081 TTTCTTTCTTAAGCCCTTGTCTCGGAGGAGTTCTCCAGCTTCAGCTCACTCAGACT 7140
7489 TCTCCAAAGCATCACCTCGGAGTTTCTCGAGGTTTCTCATAAATGAGGCTGCACATT 7548
7141 TCTCCAAAGCATCACCTCGGAGTTTCTCGAGGTTTCTCATAAATGAGGCTGCACATT 7200
7549 GCCTGTTCTGCTCGAGTATTAATACCGTCTAGTATTTTAAATGAGTGTCTAAGA 7608
7201 GCCTGTTCTGCTCGAGTATTAATACCGTCTAGTATTTTAAATGAGTGTCTAAGA 7260
7609 TTTGGTTTGGGATCAATAGGAAACATATGAGGCAACCAAGATGCAAAATGTTTGAAT 7668
7261 TTTGGTTTGGGATCAATAGGAAACATATGAGGCAACCAAGATGCAAAATGTTTGAAT 7320
7669 GATATGACCAAAATTTAAGTAGGAAAGTACCCAAACACTTCTGCTTTCACTTAAGTGT 7728
7321 GATATGACCAAAATTTAAGTAGGAAAGTACCCAAACACTTCTGCTTTCACTTAAGTGT 7380
7729 CTGGCCGATATCTGTAGGAACAGCATGATCTTGTACTGTGATATTTAAATATCCA 7788
7381 CTGGCCGATATCTGTAGGAACAGCATGATCTTGTACTGTGATATTTAAATATCCA 7440
7789 CAGTACT 7795
7441 CAGTACT 7447

RESULT 11

US-10-374-979-75
; Sequence 75, Application US/10374979
; Publication No. US20030219793A1
; GENERAL INFORMATION:
; APPLICANT: John P. Carulli et al.
; TITLE OF INVENTION: THE HIGH BONE MASS GENE OF 11q13.3
; FILE REFERENCE: 032796-021
; CURRENT APPLICATION NUMBER: US/10/374,979
; CURRENT FILING DATE: 2003-03-04
; PRIOR APPLICATION NUMBER: US 09/544,398
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 09/543,771
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 09/229,319
; PRIOR FILING DATE: 1999-01-13

; PRIOR APPLICATION NUMBER: US 60/071,449
; PRIOR FILING DATE: 1998-01-13
; PRIOR APPLICATION NUMBER: US 60/105,511
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 109
; SEQ ID NO 75
; LENGTH: 7680
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-374-979-75

Query Match 95.2%; Score 7423; DB 15; Length 7680;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 7432; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

349 GAAGAGCAAGAGGCGAGGCTCAGCAATGGTTTTCAGCCCCAGTCCCGGTGGCTGTGACGTCA 408
Db 1 GAAGAGCAAGAGGCGAGGCTCAGCAATGGTTTTCAGCCCCAGTCCCGGTGGCTGTGACGTCA 60
409 AAGCAAGCCCCGTTGTTATGACAAATGGAAAAACATATCAGATAAAATCAACAGTGGAGCG 468
Db 61 AAGCAAGCCCCGTTGTTATGACAAATGGAAAAACATATCAGATAAAATCAACAGTGGAGCG 120
469 GACCTACCTAGGCAATGCGTTGGTTTGTACTTGTATGGAGAAAGCCGAGGTTTAACTG 528
Db 121 GACCTACCTAGGTAATGTGTTGGTTTGTACTTGTATGGAGAAAGCCGAGGTTTAACTG 180
529 CAAAAGTAAACCTGAAGCTGAAGAGACTTGTCTTGACAAAGTACACTGGGAACACTTACCG 588
Db 181 CAAAAGTAAACCTGAAGCTGAAGAGACTTGTCTTGACAAAGTACACTGGGAACACTTACCG 240
589 AGTGGGTGACACTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACCTGCACTCGG 648
Db 241 AGTGGGTGACACTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACCTGCACTCGG 300
649 GCCTGGGAGGAGGAGAAATGAAGTGTACCATCGCAACCGCTGCCATGAAGGGGTGAGTC 708
Db 301 GCCTGGGAGGAGGAGAAATGAAGTGTACCATCGCAACCGCTGCCATGAAGGGGTGAGTC 360
709 CTACAAGATTTGGTGTACCTCGAGGAGACCATGAGACTGGTGTGTACATGTTTAGAGTG 768
Db 361 CTACAAGATTTGGTGTACCTCGAGGAGACCATGAGACTGGTGTGTACATGTTTAGAGTG 420
769 TGTGTGCTCTGTTAATGAAAGAGGAGTGGACCTGCAAGCCCATAGCTGAGAGAGTGT 828
Db 421 TGTGTGCTCTGTTAATGAAAGAGGAGTGGACCTGCAAGCCCATAGCTGAGAGTGT 480
829 TGATCATGCTGCTGGGACTTCTTATGTGTGGGAGAAACGTTGGGAGAGCCCTACCAAG 888
Db 481 TGATCATGCTGCTGGGACTTCTTATGTGTGGGAGAAACGTTGGGAGAGCCCTACCAAG 540
889 CTGGATGATGGTAGATTGTACTTGTCTGGGAGAGGAGCGGACGCATCACTTGCACATTC 948
Db 541 CTGGATGATGGTAGATTGTACTTGTCTGGGAGAGGAGCGGACGCATCACTTGCACATTC 600
949 TAGAAATAGATCAACAGATCAGGACACCAAGGACATCTCTATAGAATTGGAGACACCTGGAG 1008
Db 601 TAGAAATAGATCAACAGATCAGGACACCAAGGACATCTCTATAGAATTGGAGACACCTGGAG 660
1009 CAAGAAGGATTAATCAGGAAACCTCTCTCAGTGTGATCTGCACAGCAACGCGCCGAGGAGA 1068
Db 661 CAAGAAGGATTAATCAGGAAACCTCTCTCAGTGTGATCTGCACAGCAACGCGCCGAGGAGA 720
1069 GTGGAAGTGTGAGAGGCAACCTCTGTGCAGACCACTCGAGCGGATCTGGCCCCCTTCAAC 1128
Db 721 GTGGAAGTGTGAGAGGCAACCTCTGTGCAGACCACTCGAGCGGATCTGGCCCCCTTCAAC 780
1129 CGATGTTCTGTGAGTGTGTTTAAACCGGAGCCTCAACCCGAGCCTCTCTCCCTATGGCA 1188
Db 781 CGATGTTCTGTGAGTGTGTTTAAACCGGAGCCTCAACCCGAGCCTCTCTCCCTATGGCA 840
1189 CTGTGTACAGACAGTGGTGTGTTTACTCTGTGGGATGCAAGTGGCTGGAACACCAAGG 1248

b 841 CTGTGTACAGACAGTGTGTGTCTACTCTCTGTGGGATGACGTGGTTGAAGACACAGG 900
 y 1249 AAATAAGCAAAATGCTTTTGCACTGTGCTGGGCAACGGAGTCTAGCTGCCAAGAGACAGCTGT 1308
 b 901 AAATAAGCAAAATGCTTTTGCACTGTGCTGGGCAACGGAGTCTAGCTGCCAAGAGACAGCTGT 960
 y 1309 RACCAGACTTACGCTGGCAACTCAAAATGGAGAGCCATGTGCTTTACCAATTCACCTACAA 1368
 b 961 AACCCAGACTTACGCTGGCAACTTAAATGGAGAGCCATGTGCTTTACCAATTCACCTACAA 1020
 y 1369 TGGCAGGACTTCTACTCTCTGCAACCAAGAGGGGAGAGGAGACATCTTTGGTGGAG 1428
 b 1021 TGGCAGGACTTCTACTCTCTGCAACCAAGAGGGGAGAGGAGACATCTTTGGTGGAG 1080
 y 1429 CACAACTTCGAATATGACGAGCAACAGAAATACCTTTCTGACAGACCACTGTTTTT 1488
 b 1081 CACAACTTCGAATATGACGAGCAACAGAAATACCTTTCTGACAGACCACTGTTTTT 1140
 y 1489 GGTTCAGACTCGAGGAGGAAATTCCAATGGTGTGCTTGCCACTTTCCCTTCCTATACAA 1548
 b 1141 GGTTCAGACTCAAGGAGGAAATTCCAATGGTGTGCTTGCCACTTTCCCTTCCTATACAA 1200
 y 1549 CAACCAATTAACCTGATGACCTCTGAGGGGAGAGACACATCAAGTGGTGG 1608
 b 1201 CAACCAATTAACCTGATGACCTCTGAGGGGAGAGACACATCAAGTGGTGG 1260
 y 1609 GACCAACACAACTATGATGCCGAGACAGAAATTTGGGTTCTGCCCTGCTGCCACGA 1668
 b 1261 GACCAACACAACTATGATGCCGAGACAGAAATTTGGGTTCTGCCCTGCTGCCACGA 1320
 y 1669 GMAATCTGCACAACTAAGAGGGTCACTGACGGGAGAGACATGAGGATTAAGCA 1728
 b 1321 GGAAATCTGCACAACTAAGAGGGTCACTGACGGGAGAGACATGAGGATTAAGCA 1380
 y 1729 GATGACATGGGTTCACATGATGAGGTGCAGTGTGTTGGGAATGCTCGTGGGAATGGAC 1788
 b 1381 GATGACATGGGTTCACATGATGAGGTGCAGTGTGTTGGGAATGCTCGTGGGAATGGAC 1440
 y 1789 ATGCATGCTACTCCGAGTTCGAGATCAGTGCATGTTGATGACATCACTTACATGT 1848
 b 1441 ATGCATGCTACTCCGAGTTCGAGATCAGTGCATGTTGATGACATCACTTACATGT 1500
 y 1849 GAACGACATTCACCAAGCGTTCATGAAGGGGACATGCTGAACGTGACATGCTTGG 1908
 b 1501 GAACGACATTCACCAAGCGTTCATGAAGGGGACATGCTGAACGTGACATGCTTGG 1560
 y 1909 TCAGGTTCGGGGCAGGTGGAAGTGTGATCCGTCGACCAATGCCAGGATTCAGAGACTGG 1968
 b 1561 TCAGGTTCGGGGCAGGTGGAAGTGTGATCCGTCGACCAATGCCAGGATTCAGAGACTGG 1620
 y 1969 GACGTTTTATCAAAATGGAGATTCATGGGAGAGTATGTGCATGGTTCAGATACCAAGT 2028
 b 1621 GACGTTTTATCAAAATGGAGATTCATGGGAGAGTATGTGCATGGTTCAGATACCAAGT 1680
 y 2029 CTACTGCTATGGCCGTGGCAATGGGAGTGGCAATGCCAACTTTACAGACTTCCCAAG 2088
 b 1681 CTACTGCTATGGCCGTGGCAATGGGAGTGGCAATGCCAACTTTACAGACTTCCCAAG 1740
 y 2089 CTCAGTGTCTCTCGAAGTATTTATCACTGAGACTCCGAGTCCGACCACTCCCAACC 2148
 b 1741 CTCAGTGTCTCTCGAAGTATTTATCACTGAGACTCCGAGTCCGACCACTCCCAACC 1800
 y 2149 CATCCAGTGGAAATGCACCAAGCCATCTCACATTTCCAAAGTACATTTCTCAGTGGAGACC 2208
 b 1801 CATCCAGTGGAAATGCACCAAGCCATCTCACATTTCCAAAGTACATTTCTCAGTGGAGACC 1860
 y 2209 TAAATATCTGAGGCGGTGGAGAGAGTCACTACCAAGCCACTTAACTCTCTACAC 2268
 b 1861 TAAATATCTGAGGCGGTGGAGAGAGTCACTACCAAGGCGCACTTAACTCTCTACAC 1920
 y 2269 CATCAAGGCGCTGAAGCCTGGTGTATACGAGGGCCAGCTCATCAGCATCCAGAGTA 2328
 b 1921 CATCAAGGCGCTGAAGCCTGGTGTATACGAGGGCCAGCTCATCAGCATCCAGAGTA 1980

Qy 2329 CGGCCACCAAGAGTCACTCGCTTTGACTTCAACCACCAAGACACACACCTGTGAC 2388
 Db 1981 CGGCCACCAAGAGTCACTCGCTTTGACTTCAACCACCAAGACACACACCTGTGAC 2040
 Qy 2389 CAGCAACACCGTGACAGGAGAGACGATCCCTTTTCTCTTTGTGGCCACTTCTGAATC 2448
 Db 2041 CAGCAACACCGTGACAGGAGAGACGATCCCTTTTCTCTTTGTGGCCACTTCTGAATC 2100
 Qy 2449 TGTGACCGAAATCAGACGACGATGCTTTGTGCTCTCTGGTCTCAGCTTCGACACCGT 2508
 Db 2101 TGTGACCGAAATCAGACGACGATGCTTTGTGCTCTCTGGTCTCAGCTTCGACACCGT 2160
 Qy 2509 GTCCGGAATCCCGGTGGAAATATGAGCTGAGTGAGGAGGAGATGAGCCACAGTACCTGGA 2568
 Db 2161 GTCCGGAATCCCGGTGGAAATATGAGCTGAGTGAGGAGGAGATGAGCCACAGTACCTGGA 2220
 Qy 2569 TCTTCNAGCAGACGACCTCTGTGAACATCCCTGACCTGCTCTCTGGCGGAAATACAT 2628
 Db 2221 TCTTCNAGCAGACGACCTCTGTGAACATCCCTGACCTGCTCTCTGGCGGAAATACAT 2280
 Qy 2629 TGTAAATGTCTATCAGATATCTGAGATGGGAGCAGAGTTTGATCCTGTCTACTTACA 2688
 Db 2281 TGTAAATGTCTATCAGATATCTGAGATGGGAGCAGAGTTTGATCCTGTCTACTTACA 2340
 Qy 2689 AACAAAGCCCTGATGCCCTCTGACCCGACTGTGGACCAAGTTGATGACACCTCAAT 2748
 Db 2341 AACAAAGCCCTGATGCCCTCTGACCCGACTGTGGACCAAGTTGATGACACCTCAAT 2400
 Qy 2749 TGTGTTGCTGACAGACAGCCCGAGCTCCCATCAGAGGTACAGAAATAGTCTATTTCGCC 2808
 Db 2401 TGTGTTGCTGACAGACAGCCCGAGCTCCCATCAGAGGTACAGAAATAGTCTATTTCGCC 2460
 Qy 2809 ATCAGTAGAAGGTAGCAGACAGAACTCAACCTTCTGAAAATGCAAACTCCGTCAACCT 2868
 Db 2461 ATCAGTAGAAGGTAGCAGACAGAACTCAACCTTCTGAAAATGCAAACTCCGTCAACCT 2520
 Qy 2869 CAGTGACTTCAACCTGCTTCAAGTATTAACATCACTATCTATGCTGTGGAGAAATCA 2928
 Db 2521 CAGTGACTTCAACCTGCTTCAAGTATTAACATCACTATCTATGCTGTGGAGAAATCA 2580
 Qy 2929 AGAAAGTACACCTGTTGCTATTCAACAGAAACCACTGGCACCCACGCTCAGATACAGT 2988
 Db 2581 AGAAAGTACACCTGTTGCTATTCAACAGAAACCACTGGCACCCACGCTCAGATACAGT 2640
 Qy 2989 GCCCTCTCCAGGACCTGCACTTTGTGGAAGTACAGACGCTGAAGTCAACCTCATGTG 3048
 Db 2641 GCCCTCTCCAGGACCTGCACTTTGTGGAAGTACAGACGCTGAAGTCAACCTCATGTG 2700
 Qy 3049 GACACCCCTGAGAGTGCAGTGACCCGCTACCCGCTACCGTGTGATGATCCCGTCAACCTGCC 3108
 Db 2701 GACACCCCTGAGAGTGCAGTGACCCGCTACCGTGTGATGATCCCGTCAACCTGCC 2760
 Qy 3109 TGGCAGCAGCGGCAGAGGCTGCCATCAGCAGAAACACCTTTGAGAGTCAACCGGCT 3168
 Db 2761 TGGCAGCAGCGGCAGAGGCTGCCATCAGCAGAAACACCTTTGAGAGTCAACCGGCT 2820
 Qy 3169 GTCCCTCTGGGCTCACCTATTACTTCAAAAGTCTTTGCAAGTCAAGTGGAGGAGAGCAA 3228
 Db 2821 GTCCCTCTGGGCTCACCTATTACTTCAAAAGTCTTTGCAAGTCAAGTGGAGGAGAGCAA 2880
 Qy 3229 GCCTCTGACTGCTCAACAGACAAACCAACTGGATGCTCCCACTCAACTCCAGTTTGTCAA 3288
 Db 2881 GCCTCTGACTGCTCAACAGACAAACCAACTGGATGCTCCCACTCAACTCCAGTTTGTCAA 2940
 Qy 3289 TGAACCTGATTCTACTGCTCTGTGATGAGTCCACCTCGGGCCCAAGATAACAGGATA 3348
 Db 2941 TGAACCTGATTCTACTGCTCTGTGATGAGTCCACCTCGGGCCCAAGATAACAGGATA 3000
 Qy 3349 CCGACTGACCGTGGGCTTACCCGAGAGAGACGCCAGCAGTACAAATGTGGGTCCCTC 3408
 Db 3001 CCGACTGACCGTGGGCTTACCCGAGAGAGCCAGCCAGTACAAATGTGGGTCCCTC 3060

5221 AATCAACCTTGCTCTGACAGCTCATCCGTGTTGTTGATCAGAGCTTATGTTGGCCACCA 5280
5629 ATATGAAGTGAAGTGTCTATGCTTTAAGGACACCTTTGACAGGACAGCAGCTCAGGAGT 5688
5281 ATATGAAGTGAAGTGTCTATGCTTTAAGGACACCTTTGACAGGACAGCAGCTCAGGAGT 5340
5689 TGTCAACCACTCTGGAGATGTGAGCCCAACCAAGAGGCTGTGTGACAGATGCTACTGA 5748
5341 TGTCAACCACTCTGGAGATGTGAGCCCAACCAAGAGGCTGTGTGACAGATGCTACTGA 5400
5749 GACCAACCACTCTAGCTGAGAGAACCAAGACTGAGAGATCTACCTGCCCTCCAGTTGA 5808
5401 GACCAACCACTCTAGCTGAGAGAACCAAGACTGAGAGATCTACCTGCCCTCCAGTTGA 5460
5809 TCCCGTTCCAGCAATGGCCAGACTCCCAATCCAGAGAACCAAGCTCAGAGATGTCAGAG 5868
5461 TCCCGTTCCAGCAATGGCCAGACTCCCAATCCAGAGAACCAAGCTCAGAGATGTCAGAG 5520
5869 CTACACCACTCTAGCTGAGAGAACCAAGACTGAGAGATCTACCTGCCCTCCAGTTGA 5928
5521 CTACACCACTCTAGCTGAGAGAACCAAGACTGAGAGATCTACCTGCCCTCCAGTTGA 5580
5929 TGCAATGCTCGAGCTCCCTGTTGTCATCGACGCTCCACTGCCATGATGACCACTC 5988
5581 TGCAATGCTCGAGCTCCCTGTTGTCATCGACGCTCCACTGCCATGATGACCACTC 5640
5989 CAACCTGCGCTTCTGGCCACCAACCCAAATCTCTGTGTGATCATGGCAGCGCCACG 6048
5641 CAACCTGCGCTTCTGGCCACCAACCCAAATCTCTGTGTGATCATGGCAGCGCCACG 5700
6049 TGCCAGGATTCGGCTCATCATCAAGTATCAGAGCTGGGTCTCTCCAGAGAGT 6108
5701 TGCCAGGATTCGGCTCATCATCAAGTATCAGAGCTGGGTCTCTCCAGAGAGT 5760
6109 GGTCCCTCGGCCCGCTGTTGTCACAGAGCTACTATTCTGGCTGGAACCGGGAAC 6168
5761 GGTCCCTCGGCCCGCTGTTGTCACAGAGCTACTATTCTGGCTGGAACCGGGAAC 5820
6169 CGAATATCAATTTATGCTCATTTGCCCTGGAAGATATCAGAGAGCGGCTGATGG 6228
5821 CGAATATCAATTTATGCTCATTTGCCCTGGAAGATATCAGAGAGCGGCTGATGG 5880
6229 AAGGAAAAAGACAGAGCTCTCCCACTGGTAAACCTTCCACACCCCAATCTTCATGG 6288
5881 AAGGAAAAAGACAGAGCTCTCCCACTGGTAAACCTTCCACACCCCAATCTTCATGG 5940
6289 ACCAGAGATCTGGATGTTCTCCACAGTTCAAGAGCCCTTTCGTACCCACCCCTGG 6348
5941 ACCAGAGATCTGGATGTTCTCCACAGTTCAAGAGCCCTTTCGTACCCACCCCTGG 6000
6349 GTATGACACTGGAAATGGTATTTCAGCTTCTGCGACTTCTGCTGAGCAACCCAGTGTGG 6408
6001 GTATGACACTGGAAATGGTATTTCAGCTTCTGCGACTTCTGCTGAGCAACCCAGTGTGG 6060
6409 GCAACAAATGATCTTTGAGAACTGGTTTATAGGCGGACCAACCGCCCAAGAGCCAC 6468
6061 GCAACAAATGATCTTTGAGAACTGGTTTATAGGCGGACCAACCGCCCAAGAGCCAC 6120
6469 CCCATAAGGCATAGGCGGACCAATACCCGCGGATGTAGGACAGAGCTCTCTCTCA 6528
6121 CCCATAAGGCATAGGCGGACCAATACCCGCGGATGTAGGACAGAGCTCTCTCTCA 6180
6529 GACAACTATCTATGGGCCCCATTCAGAGACTTCTGAGTACATCTTATGTCATCC 6588
6181 GACAACTATCTATGGGCCCCATTCAGAGACTTCTGAGTACATCTTATGTCATCC 6240
6589 TGTGGCACTGATGAAGAACCTTACAGTTTCAGGTTCTGGAACCTTCTACAGTGCCAC 6648
6241 TGTGGCACTGATGAAGAACCTTACAGTTTCAGGTTCTGGAACCTTCTACAGTGCCAC 6300
6649 TCTGACAGGCTCACCAGAGGTCCTCAACATCATATGAGGAGGCACTGAAGACCA 6708
6301 TCTGACAGGCTCACCAGAGGTCCTCAACATCATATGAGGAGGCACTGAAGACCA 6360

QY 6709 GCAGAGGCATTAAGTTTCGGGAAAGAGTTGTTTACCGTGGGCAACTCTGTCAACGAAGGCTT 6768
Db 6361 GCAGAGGCATTAAGTTTCGGGAAAGAGTTGTTTACCGTGGGCAACTCTGTCAACGAAGGCTT 6420
QY 6769 GAACCAACCTTACGGATGACTCGTGTCTTTGACCCCTTACACAGTGTCCCATTTATGCCGTTGG 6828
Db 6421 GAACCAACCTTACGGATGACTCGTGTCTTTGACCCCTTACACAGTGTCCCATTTATGCCGTTGG 6480
QY 6829 AGATGAGTGGGAACGAATGTCTGAATCAGGCTTTTAACTGTTGTGCCAGTGTCTAGGCTT 6888
Db 6481 AGATGAGTGGGAACGAATGTCTGAATCAGGCTTTTAACTGTTGTGCCAGTGTCTAGGCTT 6540
QY 6889 TGAAGTGTGTCATTTTCAGATGATTTCTATCTAGATGTTGCCATGACAAATGTTGTGAACTA 6948
Db 6541 TGAAGTGTGTCATTTTCAGATGATTTCTATCTAGATGTTGCCATGACAAATGTTGTGAACTA 6600
QY 6949 CAAGATTGGAGAGAGTGGACCGTTCAGGAGAAATGAGGCAAGATGATGAGCTGCACATG 7008
Db 6601 CAAGATTGGAGAGAGTGGACCGTTCAGGAGAAATGAGGCAAGATGATGAGCTGCACATG 6660
QY 7009 TCTTGGAAACGGAAGAGAAATTCAGATGTTGACCCCTCATGAGGCAACCTGTTTCAGATGA 7068
Db 6661 TCTTGGAAACGGAAGAGAAATTCAGATGTTGACCCCTCATGAGGCAACCTGTTTCAGATGA 6720
QY 7069 TGGGAAGACATACCACTAGGAGAACAGTGGCAGAGGAATATCTCGGTGCCATTTGCTC 7128
Db 6721 TGGGAAGACATACCACTAGGAGAACAGTGGCAGAGGAATATCTCGGTGCCATTTGCTC 6780
QY 7129 CTGCACATGCTTTTGGAGGCCAGCGGGCTGTGACAACTGCGCGAGACCTTGGGGG 7188
Db 6781 CTGCACATGCTTTTGGAGGCCAGCGGGCTGTGACAACTGCGCGAGACCTTGGGGG 6840
QY 7189 TGAACCCAGTCCCGAGAGGCACTACTGCGCAGTCTTACAAACAGTATTTCTCAGAGATACCA 7248
Db 6841 TGAACCCAGTCCCGAGAGGCACTACTGCGCAGTCTTACAAACAGTATTTCTCAGAGATACCA 6900
QY 7249 TCAGAGAAACAAACACTAAATGTTTAAATTTGCCCAATTTAGTGTCTTCATCCCTTTAGATGTACA 7308
Db 6901 TCAGAGAAACAAACACTAAATGTTTAAATTTGCCCAATTTAGTGTCTTCATCCCTTTAGATGTACA 6960
QY 7309 GGCTGACAGAGAGATTCGCGAGAGTAAATCATCTTTCCAAATCCAGAGGAAACAGCATGT 7368
Db 6961 GGCTGACAGAGAGATTCGCGAGAGTAAATCATCTTTCCAAATCCAGAGGAAACAGCATGT 7020
QY 7369 CTCTCTGCCAAGATCCATCTAAACTGAGTGTGTTAGCAGACCCAGCTTAGAGTTCTTC 7428
Db 7021 CTCTCTGCCAAGATCCATCTAAACTGAGTGTGTTAGCAGACCCAGCTTAGAGTTCTTC 7080
QY 7429 TTTCTTTCTTAAGCCCTTTGCTCTGAGGAGTCTCCAGCTTCAGCTCACTCAGCT 7488
Db 7081 TTTCTTTCTTAAGCCCTTTGCTCTGAGGAGTCTCCAGCTTCAGCTCACTCAGCT 7140
QY 7489 TCTCAAGCATCACTTCGGAGTTCCTGAGGTTTCTCATAAATGAGGCTGCACAT 7548
Db 7141 TCTCAAGCATCACTTCGGAGTTCCTGAGGTTTCTCATAAATGAGGCTGCACAT 7200
QY 7549 GCCTGTTCTGCTCGAAGTATTCAAATACCGCTCAGTATTTTAAATGAAGTATTTAAGA 7608
Db 7201 GCCTGTTCTGCTCGAAGTATTCAAATACCGCTCAGTATTTTAAATGAAGTATTTAAGA 7260
QY 7609 TTTGTTTGGGATCAATAGGAAAGCATATGAGCAACCAAGATGCAAAATGTTTGAAT 7668
Db 7261 TTTGTTTGGGATCAATAGGAAAGCATATGAGCAACCAAGATGCAAAATGTTTGAAT 7320
QY 7669 GATATGACCAAAATTTTAAAGTAGGAAAGTCCACCAACACTTCTGCTTTCACCTAAGTGT 7728
Db 7321 GATATGACCAAAATTTTAAAGTAGGAAAGTCCACCAACACTTCTGCTTTCACCTAAGTGT 7380
QY 7729 CTGGCCCGCATATCTGAGGACAGCATGATCTGTTACTGTGATATTTTAAATATCCA 7788
Db 7381 CTGGCCCGCATATCTGAGGACAGCATGATCTGTTACTGTGATATTTTAAATATCCA 7440

2y	7789	CAGTACT 7795	1009	CAAGAAAGGATATCGAGGAAACCTGCTCCAGTGCATCTGACAGGCAACGCGGAGGAGA	1068
Db	7441	CAGTACT 7447	661	CAAGAAAGGATATCGAGGAAACCTGCTCCAGTGCATCTGACAGGCAACGCGGAGGAGA	720
RESULT 12					
US-10-447-161-4					
; Sequence 4, Application US/10447161					
; Publication No. US20040023314A1					
; GENERAL INFORMATION:					
; APPLICANT: Wang, Rong-fu					
; TITLE OF INVENTION: Mutant Fibronectin and Tumor Metastasis					
; FILE REFERENCE: HO-P0248US1					
; CURRENT APPLICATION NUMBER: US/10/447,161					
; PRIOR FILING DATE: 2003-05-28					
; PRIOR APPLICATION NUMBER: 60/383,530					
; PRIOR FILING DATE: 2002-05-28					
; NUMBER OF SEQ ID NOS: 148					
; SOFTWARE: Patent in version 3.1					
; SEQ ID NO 4					
; LENGTH: 7705					
; TYPE: DNA					
; ORGANISM: Human					
US-10-447-161-4					
Query Match 95.2%; Score 7418.2; DB 16; Length 7705;					
Best Local Similarity 99.8%; Pred. No. 0;					
Matches 7429; Conservative 0; Mismatches 18; Indels 0; Gaps 0;					
Qy	349	GAAGAGCAAGAGGAGGAGGCTCAGCAAAATGGTTTACGCCCGAGTCCCGGCTGTCAGTCA	408	1069	GTGGAAGTGTGAGAGGCACACCTCTGTGCAGACACACATCGAGCGGATCTGCCCCCTTCAC
Db	1	GAAGAGCAAGAGGAGGAGGCTCAGCAAAATGGTTTACGCCCGAGTCCCGGCTGTCAGTCA	60	721	GTGGAAGTGTGAGAGGCACACCTCTGTGCAGACACACATCGAGCGGATCTGCCCCCTTCAC
Qy	409	AAGCAAGCCCGTTTGTATGACAAATGGAACACATATCAGATTAATCAACAGTGGGAGCG	468	1129	CGATGTTGCGTGCAGCTGTTTACCAACCGCAGGCTCACCCCGAGGCTCTCTCCCTATGGCCA
Db	61	AAGCAAGCCCGTTTGTATGACAAATGGAACACATATCAGATTAATCAACAGTGGGAGCG	120	841	CTGTGTACAGACAGTGTGTGGTCTACTCTGTGGGATGAGTGTGAGAGACACAGG
Qy	469	GACCTTACCTAGCGCAATGGTTTGTACTTGTATGAGAGAGCGGAGGTTTAACTG	528	1249	AAATAAGCAAAATGCTTTGACGTGCTGGGCAACGAGTCTGCTCCCAAGAGACAGCTGT
Db	121	GACCTTACCTAGCGCAATGGTTTGTACTTGTATGAGAGAGCGGAGGTTTAACTG	180	901	AAATAAGCAAAATGCTTTGACGTGCTGGGCAACGAGTCTGCTCCCAAGAGACAGCTGT
Qy	529	CGAAGTAACTGAGCTGAGAGACTGTTTGAACAGTACACTGGGACACTTACCG	588	1309	AACCCAGACTTACGGTGGCACTCAAAATGGAGAGCATGTGTCTTACCATTCACTCAAA
Db	181	CGAAGTAACTGAGCTGAGAGACTGTTTGAACAGTACACTGGGACACTTACCG	240	961	AACCCAGACTTACGGTGGCACTTAAATGGAGAGCATGTGTCTTACCATTCACTCAAA
Qy	589	AGTGGGTGACACTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACTGCATCGG	648	1369	TGCGAGGAGCTTCTACTCTGCAACACAGAGGGCGAGGACGACATCTTTTGTGTGAG
Db	241	AGTGGGTGACACTTATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACTGCATCGG	300	1021	TGCGAGGAGCTTCTACTCTGCAACACAGAGGGCGAGGACGACATCTTTTGTGTGAG
Qy	649	GGCTGGCGAGGAGATTAAGCTGTACCATCGCAAAACCGCTGCCATGAGAGGGGTGATC	708	1429	CACAACTTCAATATGAGCAGGACGAGGACGAGGACGAGGACGAGGACGAGGACGAGG
Db	301	GGCTGGCGAGGAGATTAAGCTGTACCATCGCAAAACCGCTGCCATGAGAGGGGTGATC	360	1081	CACAACTTCAATATGAGCAGGACGAGGACGAGGACGAGGACGAGGACGAGGACGAGG
Qy	709	CTACAAGATTGGTGACACTCGAGGAGACCAATGAGACTGGTGGTTACATGTTAGAGTG	768	1489	GGTTTCAAGTCTGAGGAGGAAATTCGAATGGTGGCTTGTGCCACTTCCCCTTCTTACAA
Db	361	CTACAAGATTGGTGACACTCGAGGAGACCAATGAGACTGGTGGTTACATGTTAGAGTG	420	1141	GGTTTCAAGTCTGAGGAGGAAATTCGAATGGTGGCTTGTGCCACTTCCCCTTCTTACAA
Qy	769	TGTGTGTTTGGTAATGAAAGGAGATGAGACTTGAAGCCCATAGCTGAGAGAGTGT	828	1549	CAACCAAAATACATGATGACATCTGAGGCGGAGAGAGACACATGAAAGTGTGTGG
Db	421	TGTGTGTTTGGTAATGAAAGGAGATGAGACTTGAAGCCCATAGCTGAGAGAGTGT	480	1201	CAACCAAAATACATGATGACATCTGAGGCGGAGAGAGACACATGAAAGTGTGTGG
Qy	829	TCATCATGCTGCGGACTTCTATGCTGCGGAGAAACGTTGGAGAGCCCTACCAAGG	888	1609	GACCAACAGAACTATGATGCGGACGAGGATTTGGGTTTCTGCCCATGGCTGCCCAAGA
Db	481	TCATCATGCTGCGGACTTCTATGCTGCGGAGAAACGTTGGAGAGCCCTACCAAGG	540	1261	GACCAACAGAACTATGATGCGGACGAGGATTTGGGTTTCTGCCCATGGCTGCCCAAGA
Qy	889	CTGGATGATGTTGATTTGCTTGTGCTGGGAGAGGACGAGGACGAGGACGAGGACGAGG	948	1669	GGAAATCTGCAACCAATGAAAGGGGTCTATGATCCGATTTGGAGATCAGTGGGATGA
Db	541	CTGGATGATGTTGATTTGCTTGTGCTGGGAGAGGACGAGGACGAGGACGAGGACGAGG	600	1321	GGAAATCTGCAACCAATGAAAGGGGTCTATGATCCGATTTGGAGATCAGTGGGATGA
Qy	949	TAGAAATAGATGCAACGATCAGGACACAGGACATCTTATAGATTTGAGACACCTGGAG	1008	1729	GCAATGACATGGGTGACATGATGAGTGCACGTGTGTTGGGAATGCTCGTGGGGAATGGAC
Db	601	TAGAAATAGATGCAACGATCAGGACACAGGACATCTTATAGATTTGAGACACCTGGAG	660	1381	GCAATGACATGGGTGACATGATGAGTGCACGTGTGTTGGGAATGCTCGTGGGGAATGGAC

1741	CTCAAGTGGTCTCTGCGAAGATTTATCACTGAGACTCCGAGTCAGCCCAACTCCCACCC	1800	2821	GTCCCTGGGGTCACTTATTCTTCAAGTCTTTTGCAGTGAGCCATGGGAGGAGCAA	2880
1742	CTCAAGTGGTCTCTGCGAAGATTTATCACTGAGACTCCGAGTCAGCCCAACTCCCACCC	1801	3229	GGCTCTGACTGTCTCAACAGACAAACAACTGGATGCTCCCACTACCTCCAGTTGTCTCA	3288
1743	CATCCAGTGGAAATGACACACAGCCATCTCAATTTTCAAGTACATTTCTCAGGTGGAGAC	1802	2881	GGCTCTGACTGTCTCAACAGACAAACAACTGGATGCTCCCACTACCTCCAGTTGTCTCA	2940
1744	CATCCAGTGGAAATGACACACAGCCATCTCAATTTTCAAGTACATTTCTCAGGTGGAGAC	1803	3289	TGAAACTGATTTCTACTGTCTCTGGTGGAGTGAAGTCCACTCGGGCCCGAGATAACAGGATA	3348
1745	TAAAAATCTGTAGCCGCTTGGAGGAGTACACATACAGGCCCACTTAACTCCTACAC	1804	2941	TGAAACTGATTTCTACTGTCTCTGGTGGAGTGAAGTCCACTCGGGCCCGAGATAACAGGATA	3000
1746	TAAAAATCTGTAGCCGCTTGGAGGAGTACACATACAGGCCCACTTAACTCCTACAC	1805	3349	CGGACTGACCGTGGGCTTTACCCGAAGAGGACAGCCAGGAGTACAAATGTGGTCCCTC	3408
1747	CATCAAGGCCCTGAAGCCTGTGTGTGTATACGAGGCCAGCTCATCAGATCCAGCAGTA	1806	3001	CGGACTGACCGTGGGCTTTACCCGAAGAGGACAGCCAGGAGTACAAATGTGGTCCCTC	3060
1748	CATCAAGGCCCTGAAGCCTGTGTGTGTATACGAGGCCAGCTCATCAGATCCAGCAGTA	1807	3409	TGCTCTCAAGTACCCCACTGAGGAATCTGCAGCCTGTGATCTGAGTACACCGTATCCCTGT	3468
1749	CGGCCACCAAGAGTGAATGCTGCTTTGACTTCAACACACAGCCAGCCAGCAGCTGTGAC	1808	3061	TGCTCTCAAGTACCCCACTGAGGAATCTGCAGCCTGTGATCTGAGTACACCGTATCCCTGT	3120
1750	CGGCCACCAAGAGTGAATGCTGCTTTGACTTCAACACACAGCCAGCCAGCAGCTGTGAC	1809	3469	GGCCATAAAGGGCAACCAAGAGAGGCCCAAGGCCAATGGAGTCTTTACCACTGCAGGC	3528
1751	CAGCAACACCGTGAAGGAGAGCACTCCCTTTTCTCTCTGTGGCCACTTCTGAAATC	1810	3121	GGCCATAAAGGGCAACCAAGAGAGGCCCAAGGCCAATGGAGTCTTTACCACTGCAGGC	3180
1752	CAGCAACACCGTGAAGGAGAGCACTCCCTTTTCTCTCTGTGGCCACTTCTGAAATC	1811	3529	TGGGAGCTCTATTCCACCTTACAAACCCAGAGTGAAGTGAAGCACCATTGTGATCACATG	3588
1753	CAGCAACACCGTGAAGGAGAGCACTCCCTTTTCTCTCTGTGGCCACTTCTGAAATC	1812	3181	TGGGAGCTCTATTCCACCTTACAAACCCAGAGTGAAGTGAAGCACCATTGTGATCACATG	3240
1754	TGTGACCGAAATCAACAGCAGTAGCTTTGTGTCTCTCTGGTCTCAGCTTCGACACCGT	1813	3589	GACGCTGTCTCAAGAAATTTGGTTTAAAGTGGGTGTACGACCAAGCCAGGAGGAGGAGC	3648
1755	TGTGACCGAAATCAACAGCAGTAGCTTTGTGTCTCTCTGGTCTCAGCTTCGACACCGT	1814	3241	GACGCTGTCTCAAGAAATTTGGTTTAAAGTGGGTGTACGACCAAGCCAGGAGGAGGAGC	3300
1756	TGTGACCGAAATCAACAGCAGTAGCTTTGTGTCTCTCTGGTCTCAGCTTCGACACCGT	1815	3649	ACCACGAGAGTCACTTCAGACTCAGGAAGCATCGTTGTGTCCGGCTTGACTCCAGGAGT	3708
1757	TGTGACCGAAATCAACAGCAGTAGCTTTGTGTCTCTCTGGTCTCAGCTTCGACACCGT	1816	3301	ACCACGAGAGTCACTTCAGACTCAGGAAGCATCGTTGTGTCCGGCTTGACTCCAGGAGT	3360
1758	TGTGACCGAAATCAACAGCAGTAGCTTTGTGTCTCTCTGGTCTCAGCTTCGACACCGT	1817	3709	AGAAATAGCTTACACCATCCAGTCTCTGGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGT	3768
1759	TGTGACCGAAATCAACAGCAGTAGCTTTGTGTCTCTCTGGTCTCAGCTTCGACACCGT	1818	3361	AGAAATAGCTTACACCATCCAGTCTCTGGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGT	3420
1760	TGTGACCGAAATCAACAGCAGTAGCTTTGTGTCTCTCTGGTCTCAGCTTCGACACCGT	1819	3769	AAACAAAGTGGTGAACACCATTTGTCTCCACCAACAACTTGCATCTGGAGGCAAAACCTGA	3828
1761	TGTGACCGAAATCAACAGCAGTAGCTTTGTGTCTCTCTGGTCTCAGCTTCGACACCGT	1820	3421	AAACAAAGTGGTGAACACCATTTGTCTCCACCAACAACTTGCATCTGGAGGCAAAACCTGA	3480
1762	TGTGACCGAAATCAACAGCAGTAGCTTTGTGTCTCTCTGGTCTCAGCTTCGACACCGT	1821	3829	CACCTGAGTCTCTCAGAGTCTCTCTGGAGAGGAGCACCACCCAGACATTTACTGGTTATAG	3888
1763	TGTGACCGAAATCAACAGCAGTAGCTTTGTGTCTCTCTGGTCTCAGCTTCGACACCGT	1822	3481	CACCTGAGTCTCTCAGAGTCTCTCTGGAGAGGAGCACCACCCAGACATTTACTGGTTATAG	3540
1764	TGTGACCGAAATCAACAGCAGTAGCTTTGTGTCTCTCTGGTCTCAGCTTCGACACCGT	1823	3889	AATTAACCAACCCCTACAAAACCCGAGCAGGAGAAATTTCTTTGGAAGAAAGTGGTCCATGC	3948
1765	TGTGACCGAAATCAACAGCAGTAGCTTTGTGTCTCTCTGGTCTCAGCTTCGACACCGT	1824	3541	AATTAACCAACCCCTACAAAACCCGAGCAGGAGAAATTTCTTTGGAAGAAAGTGGTCCATGC	3600
1766	TGTGACCGAAATCAACAGCAGTAGCTTTGTGTCTCTCTGGTCTCAGCTTCGACACCGT	1825	3949	TGATCAGAGTCTCTGCACTTTTGTATTAACCTGAGTCCCGGCTGGAGTCAAAATGTCTAGTGT	4008
1767	TGTGACCGAAATCAACAGCAGTAGCTTTGTGTCTCTCTGGTCTCAGCTTCGACACCGT	1826	3601	TGATCAGAGTCTCTGCACTTTTGTATTAACCTGAGTCCCGGCTGGAGTCAAAATGTCTAGTGT	3660
1768	TGTGACCGAAATCAACAGCAGTAGCTTTGTGTCTCTCTGGTCTCAGCTTCGACACCGT	1827	4009	TTACACTGTCAAGATGAAGAAAGTGTCCCTATCTCTGTATACCATCATCCAGCTGT	4068
1769	TGTGACCGAAATCAACAGCAGTAGCTTTGTGTCTCTCTGGTCTCAGCTTCGACACCGT	1828	3661	TTACACTGTCAAGATGAAGAAAGTGTCCCTATCTCTGTATACCATCATCCAGCTGT	3720
1770	TGTGACCGAAATCAACAGCAGTAGCTTTGTGTCTCTCTGGTCTCAGCTTCGACACCGT	1829	4069	TCCTCTCCCACTGACCTGGATTCAACCAATTTGTCCAGACACCATGCGTGCACCTG	4128
1771	TGTGACCGAAATCAACAGCAGTAGCTTTGTGTCTCTCTGGTCTCAGCTTCGACACCGT	1830	3721	TCCTCTCCCACTGACCTGGATTCAACCAATTTGTCCAGACACCATGCGTGCACCTG	3780
1772	TGTGACCGAAATCAACAGCAGTAGCTTTGTGTCTCTCTGGTCTCAGCTTCGACACCGT	1831	4129	GGCTCCACCCCACTCCATTTGATTTAAACAACTTCTCTGGTGGTGTCTACCTGTGAAAAA	4188
1773	TGTGACCGAAATCAACAGCAGTAGCTTTGTGTCTCTCTGGTCTCAGCTTCGACACCGT	1832	3781	GGCTCCACCCCACTCCATTTGATTTAAACAACTTCTCTGGTGGTGTCTACCTGTGAAAAA	3840
1774	TGTGACCGAAATCAACAGCAGTAGCTTTGTGTCTCTCTGGTCTCAGCTTCGACACCGT	1833	4189	TGAGGAAGATTTGACAGTGTGCAATTTCTCTCAGACAAATGAGTGGTCTTAACAAA	4248
1775	TGTGACCGAAATCAACAGCAGTAGCTTTGTGTCTCTCTGGTCTCAGCTTCGACACCGT	1834	3841	TGAGGAAGATTTGACAGTGTGCAATTTCTCTCAGACAAATGAGTGGTCTTAACAAA	3900
1776	TGTGACCGAAATCAACAGCAGTAGCTTTGTGTCTCTCTGGTCTCAGCTTCGACACCGT	1835	4249	TCTCTCTCTCTCAGAAATATGTAGTGTGTCTCAGTGTCTACGAAACAAATGAGAG	4308
1777	TGTGACCGAAATCAACAGCAGTAGCTTTGTGTCTCTCTGGTCTCAGCTTCGACACCGT	1836	3901	TCTCTCTCTCTCAGAAATATGTAGTGTGTCTCAGTGTGTCTACGAAACAAATGAGAG	3960

3309	3361	3369	3401	3429	3481	3489	3491	3499	3509	3519	3529	3539	3549	3559	3569	3579	3589	3599	3609	3619	3629	3639	3649	3659	3669	3679	3689	3699	3709	3719	3729	3739	3749	3759	3769	3779	3789	3799	3809	3819	3829	3839	3849	3859	3869	3879	3889	3899	3909	3919	3929	3939	3949	3959	3969	3979	3989	3999	4009	4019	4029	4039	4049	4059	4069	4079	4089	4099	4109	4119	4129	4139	4149	4159	4169	4179	4189	4199	4209	4219	4229	4239	4249	4259	4269	4279	4289	4299	4309	4319	4329	4339	4349	4359	4369	4379	4389	4399	4409	4419	4429	4439	4449	4459	4469	4479	4489	4499	4509	4519	4529	4539	4549	4559	4569	4579	4589	4599	4609	4619	4629	4639	4649	4659	4669	4679	4689	4699	4709	4719	4729	4739	4749	4759	4769	4779	4789	4799	4809	4819	4829	4839	4849	4859	4869	4879	4889	4899	4909	4919	4929	4939	4949	4959	4969	4979	4989	4999	5009	5019	5029	5039	5049	5059	5069	5079	5089	5099	5109	5119	5129	5139	5149	5159	5169	5179	5189	5199	5209	5219	5229	5239	5249	5259	5269	5279	5289	5299	5309	5319	5329	5339	5349	5359	5369	5379	5389	5399	5409	5419	5429	5439	5449	5459	5469	5479	5489	5499	5509	5519	5529	5539	5549	5559	5569	5579	5589	5599	5609	5619	5629	5639	5649	5659	5669	5679	5689	5699	5709	5719	5729	5739	5749	5759	5769	5779	5789	5799	5809	5819	5829	5839	5849	5859	5869	5879	5889	5899	5909	5919	5929	5939	5949	5959	5969	5979	5989	5999	6009	6019	6029	6039	6049	6059	6069	6079	6089	6099	6109	6119	6129	6139	6149	6159	6169	6179	6189	6199	6209	6219	6229	6239	6249	6259	6269	6279	6289	6299	6309	6319	6329	6339	6349	6359	6369	6379	6389	6399	6409	6419	6429	6439	6449	6459	6469	6479	6489	6499	6509	6519	6529	6539	6549	6559	6569	6579	6589	6599	6609	6619	6629	6639	6649	6659	6669	6679	6689	6699	6709	6719	6729	6739	6749	6759	6769	6779	6789	6799	6809	6819	6829	6839	6849	6859	6869	6879	6889	6899	6909	6919	6929	6939	6949	6959	6969	6979	6989	6999	7009	7019	7029	7039	7049	7059	7069	7079	7089	7099	7109	7119	7129	7139	7149	7159	7169	7179	7189	7199	7209	7219	7229	7239	7249	7259	7269	7279	7289	7299	7309	7319	7329	7339	7349	7359	7369	7379	7389	7399	7409	7419	7429	7439	7449	7459	7469	7479	7489	7499	7509	7519	7529	7539	7549	7559	7569	7579	7589	7599	7609	7619	7629	7639	7649	7659	7669	7679	7689	7699	7709	7719	7729	7739	7749	7759	7769	7779	7789	7799	7809	7819	7829	7839	7849	7859	7869	7879	7889	7899	7909	7919	7929	7939	7949	7959
------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

5389	TGATATGAGAGCAGGCCCTGTGATGGAAACCGAGTCCACAGCTATTCTCTGCACCAACTGA	5448
5041	TGATATGAGAGCAGGCCCTGTGATGGAAACCGAGTCCACAGCTATTCTCTGCACCAACTGA	5100
5449	CGTGAAGTTCACTCAGGTACACCCACAAGCCTGAGCGCCAGTGGACACCAACCCAAATGT	5508
5101	CGTGAAGTTCACTCAGGTACACCCACAAGCCTGAGCGCCAGTGGACACCAACCCAAATGT	5160
5509	TCAGTCTCATCTGGATATCGAGTGGGGTGACCCCAAGAGAAAGACCGGACCAATGAAAGA	5568
5161	TCAGTCTCATCTGGATATCGAGTGGGGTGACCCCAAGAGAAAGACCGGACCAATGAAAGA	5220
5569	AATCAACCTTGCTCCTGACAGCTCATCCGTGTGTTATCAGGACTTATGTTGGCCACCAA	5628
5221	AATCAACCTTGCTCCTGACAGCTCATCCGTGTGTTATCAGGACTTATGTTGGCCACCAA	5280
5629	ATATGAAGTAGTGTCTATGCTCTTAAGACACTTTTGACAGCAGACACAGCTCAGGGAGT	5688
5281	ATATGAAGTAGTGTCTATGCTCTTAAGACACTTTTGACAGCAGACACAGCTCAGGGAGT	5340
5689	TGTCACCACTCTGGAGAAATGTCAGCCCAACCAAGAAGGGCTCGTGTGACAGATGCTACTGA	5748
5341	TGTCACCACTCTGGAGAAATGTCAGCCCAACCAAGAAGGGCTCGTGTGACAGATGCTACTGA	5400
5749	GACCAACATCACATTAGCTTGGAGAACCCAGACTGAGAGCATCTGCGCTCCAGTTGA	5808
5401	GACCAACATCACATTAGCTTGGAGAACCCAGACTGAGAGCATCTGCGCTCCAGTTGA	5460
5809	TGCCGTTCCAGCCAAATGGCCAGACTCCAATCCAGAGAACCATTCAAGCCAGATGTCAGAAG	5868
5461	TGCCGTTCCAGCCAAATGGCCAGACTCCAATCCAGAGAACCATTCAAGCCAGATGTCAGAAG	5520
5869	CTACACCATCACTGGCTTCAACCCAGGCACTGACTACAGAGTACTGCTGTACCTTGAA	5928
5521	CTACACCATCACTGGCTTCAACCCAGGCACTGACTACAGAGTACTGCTGTACCTTGAA	5580
5929	TGCAATGCTCGAGCTCCCTGTGTGTCATCGACGCTCCATGCGATTGATGCACCAATC	5988
5581	TGCAATGCTCGAGCTCCCTGTGTGTCATCGACGCTCCATGCGATTGATGCACCAATC	5640
5989	CAACTCGGTTTCTCGCCACCAACCCAAATCTTGTGTTATCATGCGACCGCCACG	6048
5641	CAACTCGGTTTCTCGCCACCAACCCAAATCTTGTGTTATCATGCGACCGCCACG	5700
6049	TGCCAGGATTACCGGTACATCATAGTATGAGAAGCCTGGGTCTCTCTCCAGAGAAGT	6108
5701	TGCCAGGATTACCGGTACATCATAGTATGAGAAGCCTGGGTCTCTCTCCAGAGAAGT	5760
6109	GTCCTCTCGGCCCGCCCTGGTGTCAAGAGGCTACTATTCTGCGCTTGGAAACCGGAAC	6168
5761	GTCCTCTCGGCCCGCCCTGGTGTCAAGAGGCTACTATTCTGCGCTTGGAAACCGGAAC	5820
6169	CGAATATCAATTTATGTCAATGCCCTGAAGATAATCAGAAGAGCGAGCCCTGATTGG	6228
5821	CGAATATCAATTTATGTCAATGCCCTGAAGATAATCAGAAGAGCGAGCCCTGATTGG	5880
6229	AGGAAAAGACAGACAGGCTTCCCAACTGGTAAACCTTCCACACCCCAATCTTCATGG	6288
5881	AGGAAAAGACAGACAGGCTTCCCAACTGGTAAACCTTCCACACCCCAATCTTCATGG	5940
6289	ACCAGAGATCTTGGATGTTCTTCCACAGTTCAAAAGACCCCTTTCTGTCACCCACCTGG	6348
5941	ACCAGAGATCTTGGATGTTCTTCCACAGTTCAAAAGACCCCTTTCTGTCACCCACCTGG	6000
6349	GATGACACTGGAATGTTATTCAGTTCTCTGGCATCTCTGTTAGCAACCCAGTGTGG	6408
6001	GATGACACTGGAATGTTATTCAGTTCTCTGGCATCTCTGTTAGCAACCCAGTGTGG	6060
6409	GCAACAAATGATCTTTGAGGAACATGTTTTAGCGGGAACACACCGCCCAACAGCGCCAC	6468
6061	GCAACAAATGATCTTTGAGGAACATGTTTTAGCGGGAACACACCGCCCAACAGCGCCAC	6120
6469	CCCATTAAGGCAATAGGCCACAGACCATACCCGCCGAATGTAGACAGAGAAGTCTCTCTCA	6528

1b		CCCCATTAAGGCATAGGCCAAGACCAATACCCGCCGAATGTAGGACAAAGACTCTCTCTCA	6180
1y		GACAACCATCTCATGGGCCCCATTCAGGACACTTCTCAGTACATCATTTCTGTCATCC	6588
1b		GACAACCATCTCATGGGCCCCATTCAGGACACTTCTCAGTACATCATTTCTGTCATCC	6240
1y		TGTTGGCACTGATGAGAACCCCTTACAGTTACGGGTTCCCTGGAACTTCTACCAAGTCCAC	6648
1b		TGTTGGCACTGATGAGAACCCCTTACAGTTACGGGTTCCCTGGAACTTCTACCAAGTCCAC	6300
1y		TCTGACAGGCTTACCAGAGGTGCCACCTACAACATCATAGTGGAGCACTGAAGACCA	6708
1b		TCTGACAGGCTTACCAGAGGTGCCACCTACAACATCATAGTGGAGCACTGAAGACCA	6360
1y		GCAGAGCATTAAGGTTCCGGAGAGGTTGTTACCGTGGCAACTCTCTCAACAGAGCTT	6768
1b		GCAGAGCATTAAGGTTCCGGAGAGGTTGTTACCGTGGCAACTCTCTCAACAGAGCTT	6420
1y		GAACCAACCTACGGATGACTCGTGCTTTGACCCCTACACAGTGTCCCATATGCCGTGG	6828
1b		GAACCAACCTACGGATGACTCGTGCTTTGACCCCTACACAGTGTCCCATATGCCGTGG	6480
1y		AGATGAGTGGGAACGAATGCTGAATCAGGCTTTAAACGTGTTGCCAGTGTTAGGCTT	6888
1b		AGATGAGTGGGAACGAATGCTGAATCAGGCTTTAAACGTGTTGCCAGTGTTAGGCTT	6540
1y		TGGAAGTGGTCAATTCAGATGTGAATTCATCTAGATGGTGCATGACAAATGGTGTGAACATA	6948
1b		TGGAAGTGGTCAATTCAGATGTGAATTCATCTAGATGGTGCATGACAAATGGTGTGAACATA	6600
1y		CAGATTTGGAGAGAGTGGGACCGTCAGGAGAAATGGCCAGATGATGAGCTGCACATG	7008
1b		CAGATTTGGAGAGAGTGGGACCGTCAGGAGAAATGGCCAGATGATGAGCTGCACATG	6660
1y		TCTTGGAAACGGAAAGGAAATTCAGTGTGAACCTCATGAGGCAACGTGTTACGATGA	7068
1b		TCTTGGAAACGGAAAGGAAATTCAGTGTGAACCTCATGAGGCAACGTGTTACGATGA	6720
1y		TGGGAAGACATACACGTAGGAGAACAGTGGCAGAAAGAAATATCTCGGTGCCATTTGCTC	7128
1b		TGGGAAGACATACACGTAGGAGAACAGTGGCAGAAAGAAATATCTCGGTGCCATTTGCTC	6780
1y		CTGCACATGCTTTTGGAGGCCAGCGGGCTTGGCGCTGTGACAACTGCCCGACACCTGGGGG	7188
1b		CTGCACATGCTTTTGGAGGCCAGCGGGCTTGGCGCTGTGACAACTGCCCGACACCTGGGGG	6840
1y		TGAACCCAGTCCCGAAGGCACTACTGGCCAGTCTTACAAACAGATATCTCAGAGATACCA	7248
1b		TGAACCCAGTCCCGAAGGCACTACTGGCCAGTCTTACAAACAGATATCTCAGAGATACCA	6900
1y		TCAGAGAACAAACATAATGTTAAATTTGCCAATTCAGTGTCTCATGCTTTAGATGTACA	7308
1b		TCAGAGAACAAACATAATGTTAAATTTGCCAATTCAGTGTCTCATGCTTTAGATGTACA	6960
1y		GGCTGACAGAGAGATTCCTCCGAGTAAATCATCTTTCCAAATCAGAGAAACAAGCATGT	7368
1b		GGCTGACAGAGAGATTCCTCCGAGTAAATCATCTTTCCAAATCAGAGAAACAAGCATGT	7020
1y		CTCTCTGCGAAGATCCAACTCGAGGTGATGTTAGCAGACCCAGCTTAGAGTCTTTC	7428
1b		CTCTCTGCGAAGATCCAACTCGAGGTGATGTTAGCAGACCCAGCTTAGAGTCTTTC	7080
1y		TTTCTTTCTTTAAGCCCTTTGCTCTGGAGGAAGTTCTCAGGCTTCAGCTCAACTCACAGCT	7488
1b		TTTCTTTCTTTAAGCCCTTTGCTCTGGAGGAAGTTCTCAGGCTTCAGCTCAACTCACAGCT	7140
1y		TCTCGAAGCATCACCTTGGAGTTTCTGAGGGTTTTCTCATAAATGAGGGCTGCACATT	7548
1b		TCTCGAAGCATCACCTTGGAGTTTCTGAGGGTTTTCTCATAAATGAGGGCTGCACATT	7200
1y		GCCTGTTCTGCTTCCGAAGTATTCAAATACCGCTCAGTATTTTAAATGAAGTCAATTTCAAGA	7608

Db	7201	GCCTGTTCTGCTTCGAAGTATTTCAATACCGCTCAGTATTTTAAATGAAGTGATTTCTAAGA	7256
Qy	7609	TTTGGTTTGGGATCAATAGAAAGCATATGAGCCCAACCAAGATGCAAAATGTTTTCGAAT	7668
Db	7261	TTTGGTTTGGGATCAATAGAAAGCATATGAGCCCAACCAAGATGCAAAATGTTTTCGAAT	7320
Qy	7669	GATATGACCAAAATTTTAAAGTAGAAAGTACCCAAACACTTCTGCTTTTCACTTAAAGTGT	7728
Db	7321	GATATGACCAAAATTTTAAAGTAGAAAGTACCCAAACACTTCTGCTTTTCACTTAAAGTGT	7380
Qy	7729	CTGGCCGCGCAATCTGTAGGAACAAGCATGATCTTGTACTGTGATATTTTAAATATCCA	7788
Db	7381	CTGGCCGCGCAATCTGTAGGAACAAGCATGATCTTGTACTGTGATATTTTAAATATCCA	7440
Qy	7789	CAGTACT 7795	
Db	7441	CAGTACT 7447	
RESULT 13			
US-10-098-841-6			
; Sequence 6, Application US/10098841			
; Publication No. US20020197679A1			
; GENERAL INFORMATION:			
; APPLICANT: Tang, Y. Tom			
; APPLICANT: Liu, Chenghua			
; APPLICANT: Asundi, Vinod			
; APPLICANT: Xu, Chongjun			
; APPLICANT: Zhou, Ping			
; APPLICANT: Ma, Yunging			
; APPLICANT: Wang, Jian-Rui			
; APPLICANT: Zhao, Qing A.			
; APPLICANT: Ren, Feiyan			
; APPLICANT: Chen, Rui-hong			
; APPLICANT: Wang, Dunrui			
; APPLICANT: Wang, Zhiwei			
; APPLICANT: Wehrman, Tom			
; APPLICANT: Zhang, Jie			
; APPLICANT: Qian, Xiaohong B.			
; APPLICANT: Drmanac, Radoje T.			
; TITLE OF INVENTION: No. US20020197679A1el Nucleic Acids and			
; TITLE OF INVENTION: Polypeptides			
; FILE REFERENCE: 784CIP2			
; CURRENT APPLICATION NUMBER: US/10/098,841			
; CURRENT FILING DATE: 2002-03-13			
; PRIOR APPLICATION NUMBER: 09/598,042			
; PRIOR FILING DATE: 2000-06-20			
; PRIOR APPLICATION NUMBER: 09/552,317			
; PRIOR FILING DATE: 2000-04-25			
; PRIOR APPLICATION NUMBER: 09/488,725			
; PRIOR FILING DATE: 2000-01-21			
; NUMBER OF SEQ ID NOS: 331			
; SOFTWARE: pt_Fl_genes Version 1.0			
; SEQ ID NO 6			
; LENGTH: 7867			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: (363)..(7160)			
US-10-098-841-6			
Query Match 92.7%; Score 7224.8; DB 13; Length 7867;			
Best Local Similarity 96.4%; Pred. No. 0;			
Matches 7509; Conservative 0; Mismatches 7; Indels 270; Gaps 19			
Qy	10	CCGCTCTGCTGCACAGGGGGAGGAGGGAAACCCAGCGCGGAGCGGGAAGAGGGGACC	69
Db	104	CCGCTCTGCTGCACAGGGGGAGGAGGGAAACCCAGCGCGGAGCGGGAAGAGGGGACC	163
Qy	70	TGCAGCCACAACATTCTTGTCCTCTGCAATCCCTTCTGTCCTCCACCCCGTCCCTTTCCC	129
Db	164	TGAGCCACAACATTCTTGTCCTCTGCAATCCCTTCTGTCCTCCACCCCGTCCCTTTCCC	223

	Query Match	92.7%	Score 2224.8	DB 13	Length 7867
	Best Local Similarity	96.4%	Pred. No. 0		
	Matches 7509	Conservative 0	Mismatches 7	Indels 270	Gaps 1
10	CCGCGTGTGCTGCACAGGGGAGAGAGGGAAACCCACGGGCGGAGCGGGAAGAGGGGACC	69			
104	CCGCGTGTGCTGCACAGGGGAGAGAGGGAAACCCACGGGCGGAGCGGGAAGAGGGGACC	163			
70	TGCAGCCACAACTTCTCTGTGTCCTTCATCCCTTCTGTCCCTCCACCCCGTCCCTTTCCC	129			
164	TGCAGCCACAACTTCTCTGTGTCCTTCATCCCTTCTGTCCCTCCACCCCGTCCCTTTCCC	223			

QY	130	CACCTCTGCCCCCCACCTTCTTTGGAGGCACAACCCCGGAGGCATTAGAGGATTT	189
DB	224	CACCTCTGCCCCCCACCTTCTTTGGAGGCACAACCCCGGAGGCATTAGAGGATTT	283
QY	190	TTCCGCGAGGTTGGGAAGGAGCAACCTTGGTGGCACTTGCCTCCGGTCGGGCGTC	249
DB	284	TTCCGCGAGGTTGGGAAGGAGCAACCTTGGTGGCACTTGCCTCCGGTCGGGCGTC	343
QY	250	TCTCCCCACCGCTCTCAACATGCTTAGGGGTCCGGGCGCGGCTCTCTCTGTGCGCCT	309
DB	344	TCTCCCCACCGCTCTCAACATGCTTAGGGGTCCGGGCGCGGCTCTCTCTGTGCGCCT	403
QY	310	CCTGTGCTCTGGGACACGCGTGCCTTCCACGGGAGGCTCGAAGAGCAAGGCGAGCTCA	369
DB	404	CCTGTGCTCTGGGACACGCGTGCCTTCCACGGGAGGCTCGAAGAGCAAGGCGAGCTCA	463
QY	370	GCAAAATGGTTTCAGCCCCAGTCCCGGTGGTGTCAGTCAAAGCAAGCCCGTTGTTATGA	429
DB	464	GCAAAATGGTTTCAGCCCCAGTCCCGGTGGTGTCAGTCAAAGCAAGCCCGTTGTTATGA	523
QY	430	CAATGGAAACACTATCAGATTAATCAACAGTGGGAGCGACCTACTCTAGGCAATCGCTT	489
DB	524	CAATGGAAACACTATCAGATTAATCAACAGTGGGAGCGACCTACTCTAGGCAATCGCTT	583
QY	490	GGTTGTGACTTGTATGGAGGAAGCCGAGGTTTTAACTGCGAAAGTAAACCTCGAAGCTGA	549
DB	584	GGTTGTGACTTGTATGGAGGAAGCCGAGGTTTTAACTGCGAGAGTAAACCTCGAAGCTGA	643
QY	550	AGAGACTTGCTTTGACAAGTACACTGGGAACACTTACCGAGTGGGTGACACTTATGAGCG	609
DB	644	AGAGACTTGCTTTGACAAGTACACTGGGAACACTTACCGAGTGGGTGACACTTATGAGCG	703
QY	610	TCCTAAAGACTCCATGATCTGGGACTGTACTGTGATCGGGGCTGGGCGAGGAGAAATAG	669
DB	704	TCCTAAAGACTCCATGATCTGGGACTGTACTGTGATCGGGGCTGGGCGAGGAGAAATAG	763
QY	670	CTGTACCATCGCAAAACCGCTGCCATGAAGGGGGTCAGTCCTACAAGNTTGGTGACACCTG	729
DB	764	CTGTACCATCGCAAAACCGCTGCCATGAAGGGGGTCAGTCCTACAAGNTTGGTGACACCTG	823
QY	730	GAGGAGACCACATCAGACTGTGTGTTACATGTTAGAGTGTGTGTCTTGTTAAATGGAAA	789
DB	824	GAGGAGACCACATCAGACTGTGTGTTACATGTTAGAGTGTGTGTCTTGTTAAATGGAAA	883
QY	790	AGGAGAATGNACTGCAAGCCATAGCTGAGAAGTGTTTTGATCATGCTCTCTGGACTTC	849
DB	884	AGGAGAATGNACTGCAAGCCATAGCTGAGAAGTGTTTTGATCATGCTCTCTGGACTTC	943
QY	850	CTATGTGCTCGGAGAAAACGTGGGAGAAAGCCCTACCAAGGCTGGATCATGTTAGATTGTAC	909
DB	944	CTATGTGCTCGGAGAAAACGTGGGAGAAAGCCCTACCAAGGCTGGATCATGTTAGATTGTAC	1003
QY	910	TTGCTCTGGGAGAGCGAGACGCATCATCTTGACCTTTAGAAAATAGATGCAACGATCA	969
DB	1004	TTGCTCTGGGAGAGCGAGACGCATCATCTTGACCTTTAGAAAATAGATGCAACGATCA	1063
QY	970	GGACAAGAGCATCTATAGAAATTGGAGACACCTCGGAGCAAGAGATTAATCGAGAAA	1029
DB	1064	GGACAAGAGCATCTATAGAAATTGGAGACACCTCGGAGCAAGAGATTAATCGAGAAA	1123
QY	1030	CCTGCTCAGTGATCTTGCAAGGGCAACGGCCGAGGAGTGGAAAGTGTGAGAGGCACAC	1089
DB	1124	CCTGCTCAGTGATCTTGCAAGGGCAACGGCCGAGGAGTGGAAAGTGTGAGAGGCACAC	1183
QY	1090	CTCTGTGAGACCAATCGAGCGGATCTGGCCCTTACCGATCTTCTGTGAGCTGTTTA	1149
DB	1184	CTCTGTGAGACCAATCGAGCGGATCTGGCCCTTACCGATCTTCTGTGAGCTGTTTA	1243
QY	1150	CCAACCGAGCGCTCACCCCAAGCCTCTCTCCCTATGTGGCCACTGTGTCAACAGTGTGT	1209
DB	1244	CCAACCGAGCGCTCACCCCAAGCCTCTCTCCCTATGTGGCCACTGTGTCAACAGTGTGT	1303

Qy	1210	GGTCTACTCTGTGGGATGCGATGGCTGAAGACACAAGAAATTAAGCAATGCTTTGCAC	1269
Db	1304	GGTCTACTCTGTGGGATGCGATGGCTGAAGACACAAGAAATTAAGCAATGCTTTGCAC	1363
Qy	1270	GTCCCTGGGCAACGGAGTCAGCTGCCAAGAGACAGCTGTAAACCGACACTTACGGTGGCAA	1329
Db	1364	GTCCCTGGGCAACGGAGTCAGCTGCCAAGAGACAGCTGTAAACCGACACTTACGGTGGCAA	1423
Qy	1330	CTCAAAATGGAGACCCATGTGTCTTACATTCACCTACATTAATGGCAGGACGTTCTACTCTCTG	1389
Db	1424	CTCAAAATGGAGACCCATGTGTCTTACATTCACCTACATTAATGGCAGGACGTTCTACTCTCTG	1483
Qy	1390	CACCAAGAGAGGCGACAGACGGACATCTTTGGTGCAGCACAACTTCGAATATTAGCA	1449
Db	1484	CACCAAGAGAGGCGACAGACGGACATCTTTGGTGCAGCACAACTTCGAATATTAGCA	1543
Qy	1450	GGACCAAGAAATACTCTTTCTGCAACAGACCACTGTTTGGTTTCAGACTCGAGGAGGAAA	1509
Db	1544	GGACCAAGAAATACTCTTTCTGCAACAGACCACTGTTTGGTTTCAGACTCGAGGAGGAAA	1603
Qy	1510	TTCCAATGGTGCCTTGTGCACACTTCGGCTTCCTATACACAAACCAANTTACACTGATTCG	1569
Db	1604	TTCCAATGGTGCCTTGTGCACACTTCGGCTTCCTATACACAAACCAANTTACACTGATTCG	1663
Qy	1570	CACCTTCAGGGGAGAGACAACTGAAGTGGTGTGGGACCAACAGAACTATGATGC	1629
Db	1664	CACCTTCAGGGGAGAGACAACTGAAGTGGTGTGGGACCAACAGAACTATGATGC	1723
Qy	1630	CGACCAAGAAATTTGGGTTCTGCCCATGGCTGCCACAGGAAATCTGCACACCAATGA	1689
Db	1724	CGACCAAGAAATTTGGGTTCTGCCCATGGCTGCCACAGGAAATCTGCACACCAATGA	1783
Qy	1690	AGGGTCATCTACCGCATTCGGAGATCAGTGGGATAGCAGCATGAATGGGTCAACATGAT	1749
Db	1784	AGGGTCATCTACCGCATTCGGAGATCAGTGGGATAGCAGCATGAATGGGTCAACATGAT	1843
Qy	1750	GAGGTGCACGTGTGTGGGAATGGTGGTGGGAAATGGACATGTCAATTCGCTACTTCGCACT	1809
Db	1844	GAGGTGCACGTGTGTGGGAATGGTGGTGGGAAATGGACATGTCAATTCGCTACTTCGCACT	1903
Qy	1810	TCGAGATCAGTGCATTTTCATGACATCACTTACATGTCAATGTCAAGACACATTCACCAAGCG	1869
Db	1904	TCGAGATCAGTGCATTTTCATGACATCACTTACATGTCAAGACACATTCACCAAGCG	1953
Qy	1870	TCATGAAGAGGGGCACATGCTGAATCTTCATGCTTCGGTCAGGGTCGGGGCAGGTGGAA	1929
Db	1964	TCATGAAGAGGGGCACATGCTGAATCTTCATGCTTCGGTCAGGGTCGGGGCAGGTGGAA	2023
Qy	1930	GTGTGATCCCGTCGACCAATGCCAGGATTCAGAGACTGGGACGTTTTATCAAAATGGAGA	1989
Db	2024	GTGTGATCCCGTCGACCAATGCCAGGATTCAGAGACTGGGACGTTTTATCAAAATGGAGA	2083
Qy	1990	TTCAATGGGGAAGTATGTGCATGGTGTTCAGATACCAGTGTCTTCGTATATGGCCGTGGCAT	2049
Db	2084	TTCAATGGGGAAGTATGTGCATGGTGTTCAGATACCAGTGTCTTCGTATATGGCCGTGGCAT	2143
Qy	2050	TGGGGATGGCATTTGCCAAACCTTTACAGACCTATCAGCTCAAGTGGTCTCTGTGCAAGT	2109
Db	2144	TGGGGATGGCATTTGCCAAACCTTTACAGACCTATCAGCTCAAGTGGTCTCTGTGCAAGT	2203
Qy	2110	ATTTATCACTGAGACTCCGAGTTCAGCCCACTCCACCCCATCCAGTGGAAATGCACACA	2169
Db	2204	ATTTATCACTGAGACTCCGAGTTCAGCCCACTCCACCCCATCCAGTGGAAATGCACACA	2263
Qy	2170	GCCATCTCAATTTCCAAATGACATTTCTCAGGTGGAGACCTTAAAAATTTCTGTAGGCCGTTG	2239
Db	2264	GCCATCTCAATTTCCAAATGACATTTCTCAGGTGGAGACCTTAAAAATTTCTGTAGGCCGTTG	2323
Qy	2230	GAAGGAGCTACCATACAGGCCACTTAACTCCTACACCATCAAGGCCCTGAGGCCCTGG	2289
Db	2324	GAAGGAGCTACCATACAGGCCACTTAACTCCTACACCATCAAGGCCCTGAGGCCCTGG	2383
Qy	2290	TGTGGTATACGAGGGCCAGCTCATCAGCATCCAGCAGTACGGGCCACCAAGAGTACTCG	2349

QY	4510	CCTCACCACCTCACTCCAGGACAGAGTATGTGGTCAGCATCGTTCCTTAATGCGAG	4569
DB	4604	CCTCACCACCTCACTCCAGGACAGAGTATGTGGTCAGCATCGTTCCTTAATGCGAG	4663
QY	4570	AGAGGAAAGTCCCTTATTGATTGGCCCAACAATCAACAGTTTCTGATGTTCCGAGGGACCT	4629
DB	4664	AGAGGAAAGTCCCTTATTGATTGGCCCAACAATCAACAGTTTCTGATGTTCCGAGGGACCT	4723
QY	4630	GGAGTTGTTGTCGACGCCCAACAGCTACTGATGAGTGGGATGCTCTGCTGTCTCAC	4689
DB	4724	GGAGTTGTTGTCGACGCCCAACAGCTACTGATGAGTGGGATGCTCTGCTGTCTCAC	4783
QY	4690	AGTGAGATATTACAGGATCACTTACGGGAAACAGGAGGAATAAGCCCTGTCACGAGATT	4749
DB	4784	AGTGAGATATTACAGGATCACTTACGGGAAACAGGAGGAATAAGCCCTGTCACGAGATT	4843
QY	4750	CACCTGTGCTGGAGCAGTCTACAGCTACCATCAGCGGCTTAAACCTGAGTTGATTA	4809
DB	4844	CACCTGTGCTGGAGCAGTCTACAGCTACCATCAGCGGCTTAAACCTGAGTTGATTA	4903
QY	4810	TACCATCACTGTGTATGCTCACTTGGCGGTGGAGACAGCGCCCGCAAGCAGCAAGCCAAT	4869
DB	4904	TACCATCACTGTGTATGCTCACTTGGCGGTGGAGACAGCGCCCGCAAGCAGCAAGCCAAT	4963
QY	4870	TTCGATTAATACCGAAACAGAAATTGA CAACCAATCCAGATGCCAGTGCAGATGTTCA	4929
DB	4964	TTCGATTAATACCGAAACAGAAATTGA CAACCAATCCAGATGCCAGTGCAGATGTTCA	5023
QY	4930	GGACAAACAGCATTAAGTGTCAAGTGGCTGCTCCCTCAAGTTCCCTGTTACTGGTTACAGAGT	4989
DB	5024	GGACAAACAGCATTAAGTGTCAAGTGGCTGCTCCCTCAAGTTCCCTGTTACTGGTTACAGAGT	5083
QY	4990	AACCAACCACTCCCAAAAATGACAGGACCAACAAAACCTAAAACTGCAGTCCAGATCA	5049
DB	5084	AACCAACCACTCCCAAAAATGACAGGACCAACAAAACCTAAAACTGCAGTCCAGATCA	5143
QY	5050	AACAGAAATGACTATTGAAGGCTTCAGCCACAGTGGATGTTAGTGTCTATGC	5109
DB	5144	AACAGAAATGACTATTGAAGGCTTCAGCCACAGTGGATGTTAGTGTCTATGC	5203
QY	5110	TCAGAAATCCAGCGGAGAGTCAAGCTCTGGTTCCAGTCCAGTAAACCAATGATCG	5169
DB	5204	TCAGAAATCCAGCGGAGAGTCAAGCTCTGGTTCCAGTCCAGTAAACCA	5253
QY	5170	CCCTAAGGACTGGCATTTCACTGATGTGGATGTCGATTCCATCAAAATGCTTGGGAAG	5229
DB	5254	-----	5253
QY	5230	CCCACAGGGGCAAGTTTCCAGGTACAGGGTGACCTACTCGAGCCCTGAGGATGGAATCCA	5289
DB	5254	-----	5253
QY	5290	TGAGCTATTCCCTGCACCTGATGGTGAAGAAGACACTGCAGAGCTGCAGGCTCAGACC	5349
DB	5254	-----	5253
QY	5350	GGGTTCTGAGTACACAGTCAGTGTGGTTGCCCTTGCCAGATGATGAGAGCCAGCCCT	5409
DB	5254	-----	5253
QY	5410	GATTGGAACCCAGTCCACAGCTATTCTGCAACCAACTGACCTGAAGTTCACTCAGGTCAAC	5469
DB	5254	-----CTATTCTGCAACCAACTGACCTGAAGTTCACTCAGGTCAAC	5293
QY	5470	ACCCACAAAGCTGAGCGCCAGTGAGACACCACTGTCAGTCTCAGTGTATTCGAT	5529
DB	5294	ACCCACAAAGCTGAGCGCCAGTGAGACACCACTGTCAGTCTCAGTGTATTCGAT	5553
QY	5530	GGGGTGACCCCAAGGAGAACCGGACCAATGAAGAAATCAACCTTGTCTCTGACAG	5589
DB	5354	GGGGTGACCCCAAGGAGAACCGGACCAATGAAGAAATCAACCTTGTCTCTGACAG	5413

QY	5590	CTCATCCGTGGTTGTATCAGGACTTATGGTGGCCACCAAAATATGAAGTGAGTGTCTATGC	5649
DB	5414	CTCATCCGTGGTTGTATCAGGACTTATGGTGGCCACCAAAATATGAAGTGAGTGTCTATGC	5473
QY	5650	TCTTAAGGACATTTTGA CAAGCAGACCTCAGGAGTGTGACCACTCTCGAGAAATGT	5709
DB	5474	TCTTAAGGACATTTTGA CAAGCAGACCTCAGGAGTGTGACCACTCTCGAGAAATGT	5533
QY	5710	CAGCCCAACAAAGAGGCTGCTGACAGATCTACTGAGACCAACCATCACCATTAGCTG	5769
DB	5534	CAGCCCAACAAAGAGGCTGCTGACAGATCTACTGAGACCAACCATCACCATTAGCTG	5593
QY	5770	GAGAAACAAAGACTGAGACGATCACTGGGCTTCCAAAGTTGATGCCGTTCCAGCCAAATGGCCA	5829
DB	5594	GAGAAACAAAGACTGAGACGATCACTGGGCTTCCAAAGTTGATGCCGTTCCAGCCAAATGGCCA	5653
QY	5830	GACTCCAAATCCAGAGAACCATCAAGCCAGATGTGAGAAGCTACACCATCAGTGGCTTACA	5889
DB	5654	GACTCCAAATCCAGAGAACCATCAAGCCAGATGTGAGAAGCTACACCATCAGTGGCTTACA	5713
QY	5890	ACCAGGCACTGACTACAAGATCTA CTTGTACACCTTGAATGACAAATGCTCGGAGTCCCC	5949
DB	5714	ACCAGGCACTGACTACAAGATCTA CTTGTACACCTTGAATGACAAATGCTCGGAGTCCCC	5773
QY	5950	TGTGGTCACTGACGCTCCACTGCCATTCGATGACCATCCAACTCCGCTTCTCGGCCAC	6009
DB	5774	TGTGGTCACTGACGCTCCACTGCCATTCGATGACCATCCAACTCCGCTTCTCGGCCAC	5833
QY	6010	CACACCAATTTCTTGTGTTATCATGGCAGCGGCCACCGTCCAGGATTAACCGGCTACAT	6069
DB	5834	CACACCAATTTCTTGTGTTATCATGGCAGCGGCCACCGTCCAGGATTAACCGGCTACAT	5893
QY	6070	CATCAAGTATGAGAGCTGGGTCTCTCCAGAGAGTGGTCCCTCGGCCCGGCTGG	6129
DB	5894	CATCAAGTATGAGAGCTGGGTCTCTCCAGAGAGTGGTCCCTCGGCCCGGCTGG	5953
QY	6130	TGTCAAGAGGCTTACTTATTTACTTGGCTCGAAACCGGAACCGGAATATACAATTTATGTCA	6189
DB	5954	TGTCAAGAGGCTTACTTATTTACTTGGCTCGAAACCGGAACCGGAATATACAATTTATGTCA	6013
QY	6190	TGCTCTGAAGAATAATACAGAGAGCGGCTTGTATGGAAGGAAAGACAGAGAGCT	6249
DB	6014	TGCTCTGAAGAATAATACAGAGAGCGGCTTGTATGGAAGGAAAGACAGAGAGCT	6073
QY	6250	TCCCAACTGGTAAACCTTCCACACCCCAATTTCTATGGACAGAGATCTTGGATGTTCC	6309
DB	6074	TCCCAACTGGTAAACCTTCCACACCCCAATTTCTATGGACAGAGATCTTGGATGTTCC	6133
QY	6310	TTCCACAGTTCAAAAGACCCCTTTTCTGTCACCCACCTGGTATGACACTGGAAATGGTAT	6369
DB	6134	TTCCACAGTTCAAAAGACCCCTTTTCTGTCACCCACCTGGTATGACACTGGAAATGGTAT	6193
QY	6370	TCAGCTTCTGTCACCTTCTGGTCAGCAACCCAGTGTGGGCAACCAATGATCTTTGAGGA	6429
DB	6194	TCAGCTTCTGTCACCTTCTGGTCAGCAACCCAGTGTGGGCAACCAATGATCTTTGAGGA	6253
QY	6430	ACATGTTTTTAGGCGGACCAACCCGCCACCAACCGGCCACCCCCATAAGGCCATAGGCCAAG	6489
DB	6254	ACATGTTTTTAGGCGGACCAACCCGCCACCAACCGGCCACCCCCATAAGGCCATAGGCCAAG	6313
QY	6490	ACCATACCCCGGAAATGATAGGACAGAGCTCTCTCTCAGACCAACCATCTCATGGGCCCC	6549
DB	6314	ACCATACCCCGGAAATGATAGGACAGAGCTCTCTCTCAGACCAACCATCTCATGGGCCCC	6373
QY	6550	ATTCCAGGACACTTCTGAGTACATCATTTTCTGTCATCTGTTGGCACTGATGAAGAACC	6609
DB	6374	ATTCCAGGACACTTCTGAGTACATCATTTTCTGTCATCTGTTGGCACTGATGAAGAACC	6433
QY	6610	CTTAAGTTTACGGGTTCTTGGAACTTTCTACAGTGCACCTCTGACAGGCTCACCAGAGG	6669
DB	6434	CTTAAGTTTACGGGTTCTTGGAACTTTCTACAGTGCACCTCTGACAGGCTCACCAGAGG	6493
QY	6670	TGCCACCTTACATCATATAGTGGAGGCACTGAAAGACAGCAGAGGACATTAAGGTTCCGGGA	6729

6494 TGCACCTACACATCATAGTGGAGGCACTGAAGACACAGACAGGCATAGGTTCCGGA 6553
6730 AGAGGTTGTTTACCGTGGGCAACTCTGTCAACGAAGGCTTGAACCAACCTACGGATGACTC 6789
6554 AGAGGTTGTTTACCGTGGGCAACTCTGTCAACGAAGGCTTGAACCAACCTACGGATGACTC 6613
6790 GTGCTTTGACCCCTACACAGTGTCCCATTTATCCGTTGGAGTGAAGTGGGAACGATGTC 6849
6614 GTGCTTTGACCCCTACACAGTGTCCCATTTATCCGTTGGAGTGAAGTGGGAACGATGTC 6673
6850 TGAATCAGGCTTTAAACTGTTGTGCCAGTGCTTAGGCTTTGGAAGTGGTCAATTTCCAGATG 6909
6674 TGAATCAGGCTTTAAACTGTTGTGCCAGTGCTTAGGCTTTGGAAGTGGTCAATTTCCAGATG 6733
6910 TGAATCAGTGTGTCATGATGACATGATGGTGTGACATACAGTGTGACATGATGGAGAGTGGGA 6969
6734 TGAATCAGTGTGTCATGATGACATGATGGTGTGACATACAGTGTGACATGATGGAGAGTGGGA 6793
6970 CCGTCAGGAGAAAATGGCCAGATGATGAGTGTGCAATGCTTGGGAAACGGAAGAGAGA 7029
6794 CCGTCAGGAGAAAATGGCCAGATGATGAGTGTGCAATGCTTGGGAAACGGAAGAGAGA 6853
7030 ATTCAAGTGTGACCTTCATGAGGCAACGTTTACGATGATGGAAGACATACACGTTAGG 7089
6854 ATTCAAGTGTGACCTTCATGAGGCAACGTTTACGATGATGGAAGACATACACGTTAGG 6913
7090 AGAACAGTGGCAGAGAAATATCTCGGTGCCATTTGCTCTGCAATGCTTTGGAGGCCA 7149
6914 AGAACAGTGGCAGAGAAATATCTCGGTGCCATTTGCTCTGCAATGCTTTGGAGGCCA 6973
7150 GCGGGGCTGGCGTGTGACAACTGCGCGAGACTGGGGGTGAACCCAGTCCCGAAGGCAC 7209
6974 GCGGGGCTGGCGTGTGACAACTGCGCGAGACTGGGGGTGAACCCAGTCCCGAAGGCAC 7033
7210 TACTGCCAGTCTCAACACAGTATTTCTCAGATACCATCAGAGAACAAACACTAATGT 7269
7034 TACTGCCAGTCTCAACACAGTATTTCTCAGATACCATCAGAGAACAAACACTAATGT 7093
7270 TAATGCCAAATGAGTGTTCATGCTTTAGATGATACAGGTGACAGAGAGATTCGCG 7329
7094 TAATGCCAAATGAGTGTTCATGCTTTAGATGATACAGGTGACAGAGAGATTCGCG 7153
7330 AGAGTAAATCATCTTCCAAATCAGAGGAAACAGCATGCTCTCTCCCAAGATCCATGTA 7389
7154 AGAGTAAATCATCTTCCAAATCAGAGGAAACAGCATGCTCTCTCCCAAGATCCATGTA 7213
7390 AACTGGAGTGAATGATAGCAGACCCAGCTTAGAGTCTTCTTTCTTTTAAAGCCCTTTGC 7449
7214 AACTGGAGTGAATGATAGCAGACCCAGCTTAGAGTCTTCTTTCTTTTAAAGCCCTTTGC 7273
7450 TCTGGAGGAGATTTCTCAGCTTCACTCACTCAGCTTCACTCAGCTTCACTCAGCTTCACT 7509
7274 TCTGGAGGAGATTTCTCAGCTTCACTCAGCTTCACTCAGCTTCACTCAGCTTCACT 7333
7510 GTTTCCTGAGGTTTCTCATAAATGAGGCTGCACATTCGCTGTTCTGCTTGAAGTAT 7569
7334 GTTTCCTGAGGTTTCTCATAAATGAGGCTGCACATTCGCTGTTCTGCTTGAAGTAT 7393
7570 TCAATACCGTTCAGTATTTTAAATGAGTGAATTTGTTTGGATCAATAGGA 7629
7394 TCAATACCGTTCAGTATTTTAAATGAGTGAATTTGTTTGGATCAATAGGA 7453
7630 AAGCATATGACGCAACCAAGATGCAAAATGTTTGAATGATATGACCAAAATTTTAACT 7689
7454 AAGCATATGACGCAACCAAGATGCAAAATGTTTGAATGATATGACCAAAATTTTAACT 7513
7690 AGGAAGTCCACCAACACTCTGCTTCACTTAACTGCTGCGCGCAATCTGTAGGA 7749
7514 AGGAAGTCCACCAACACTCTGCTTCACTTAACTGCTGCGCGCAATCTGTAGGA 7573
7750 ACAAGCATGATCTGTGATGATATTTTAAATATCCACAGTACT 7795

Db 7574 ACAAGCATGATCTTGTACTGTGATATTTTAAATATCCACAGTACT 7619
RESULT 14
US-09-917-800A-1731
; Sequence 1731, Application US/09917800A
; Patent No. US20020119462A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Castle, Arthur
; APPLICANT: Blashoff, Michael
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5038-US
; CURRENT APPLICATION NUMBER: US/09/917,800A
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/222,040
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 60/222,880
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: US 60/290,029
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/290,645
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: US 60/292,336
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/295,798
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/297,457
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,884
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,459
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 1740
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1731
; LENGTH: 8329
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 NM_019143
US-09-917-800A-1731
Query Match 68.3%; Score 5320.6; DB 9; Length 8329;
Best Local Similarity 81.8%; Pred. No. 0; Mismatches 1044; Indels 397; Gaps 14;
Matches 6484; Conservative 0;
QY 184 GGAATTTTCCCGCAGGTTGCGAAGGGAAGCAACTTGGTGGCAACTTGCCTCCCGGTGCG 243
Db 123 GGACTTTTCTCGCAGGCTGCGAGGGAA-CCAACTTCTGGCCACTAGCCTCTCGGAGAG 181
QY 244 GGGTCTCTCC-CCACCGTCTCAACATGCTTAGGGTTCGGGGCCCG---GGCTGCTGC 299
Db 182 GGCAGCTCTCTCCCTCCACTCCAGATGCTCAGGGTCCGGGACCCGGCGGCTGCTGC 241
QY 300 TGCTGGCGCTCTCTGCTGGGACAGCGTGCCTCCACGGGAGCCTCGAAGAGCAAGA 359
Db 242 TGCTAGCAGTCTCTGCTGGGACATCGGTGCGCTGCACCGAAACCGGAAGAGCAAGA 301
QY 360 GGCAGGCTCAGCAATAGTTCAGCCCCAGTCCCGGTGGGTGCTCAGTCAAGAACGCCG 419
Db 302 GGCAGGCTCAGCAATAGTTCAGCCCCAGTCCCGGTGGGTGCTCAGTCAAGAACGCCG 361
QY 420 GTTGTATGACAAATGGAACACACTATCAGATAAATCAACAGTGGGAGCGGACCTACCTAG 479
Db 362 GCTGTTTGACAAAGGAGCAATATCAGATAATCAGCAGTGGGACGACCTACCTAG 421
QY 480 GCAATGCTTGGTTGACTTGTATGAGAACCCAGGTTTAACTTGAAGTAACTGCAAGTAAAC 539
Db 422 GCAACGCCCTGGTTGTACCTGCTATGAGGAAGCAGAGGTTTAACTGCGAGGCAAGC 481

2Y	540	CTGAAGCTGAAGAGACTTGTCTTGACAAGTACACTGGGAACACTTACCGAGTGGGTGACA	599
Db	482	CTGAACTCTGAAGAGACTGTTTTTGACAAATACACTGGAAACACTTACAAAGTGGGTGACA	541
2Y	600	CTTATGAGCGTCTTAAGACTCCATGATCTGGGACTGTACTGCTGCTATCGGGCTGGCGGAG	659
Db	542	CTTATGAGCGCTTAAGATTCATGATCTGGGACTGTACTCTGATCTGGGGCTGGCGGAG	601
2Y	660	GGAGAAATAAGCTGTACCATCGCAAAACCGCTGCCATGAAGGGGGTCACTCTCAAGATTG	719
Db	602	GCAGGATCAGCTGTACCAATTGCAATTCGCTGCCATGAAGGGGGTCACTCTCAAGATTG	661
2Y	720	GTGACACCTGGAGGAGACACATGAGACTGGTGGTTACATGTTAGAGTGTGTGTGTCCTG	779
Db	662	GTGACAAGTGGAGGAGGCCACATGAGACTGGTGGCTATATGTTGGAGTGTGTTGTGTCTGG	721
2Y	780	GTAATGAAAAGAGAGAAATGCACTCGAAGCCCATAGCTGAGAAAGTGTGTTGATCATGCTG	839
Db	722	GGAAATGAAAAGGGGAATGCACTCGAAGCCCATAGCTGAGAAATGTTTGTATCAGCTG	781
2Y	840	CTGGGACTTCTATGTGGTCTGGAGAAACGTGGGAGAGCCCTTACCAAGGCTGATGATGG	899
Db	782	CTGGGACTTCTACGTCTGGGGGAGACCTCGGAAAGCCCTTCAAGGCTGGATGATGG	841
2Y	900	TAGATTGTACTTCCTGGGAGAGGACGCGACATCACTTCCACTTCTAGAAAATAGAT	959
Db	842	TGGACTGTATCTTGTCTGGCGAAGGCNAATGGCGGTATCACCTGCACTCTCCGGAAACAGAT	901
2Y	960	GCAACGATCAGGACACAAGGACATCTCTATAGAAATGGAGACACCTGGAGCAAGAGGATA	1019
Db	902	GCAATGATCAGGACACCAAGGACGCTCTACAGAAATGGAGACACATGGAGCAAGAGGACA	961
2Y	1020	ATCGAGAAACCTGCTCCAGTGCATCTGCACAGCAACGCGCGAGGAGAGTGAAGTGTG	1079
Db	962	ACAGAGGGAACCTGCTCCAGTGTCTGCACAGCAACGCGCAGAGGGAGTGAAGTGTG	1021
2Y	1080	AGAGGCACACCTCTGTGCAGACCAACATCGAGCGGATCTGGCCCTTCCACCGATGTTCTGTG	1139
Db	1022	AGGCACA--TGTTCTACAGAGTGCCTTCACTGGAATCTGCTCTCTTACAGATGTCCGAA	1078
2Y	1140	CAGCTGTTTACGAACCGCAGCCTCACCCCGAGCCTCTCTCCCTATGSCCATGTGTACAG	1199
Db	1079	CAGCTATTTACCAACCCCGACACCCCGCAGCCCGCACCGTACGGCCACTGTGTCAAG	1138
2Y	1200	ACAGTGTGTGCTACTCTGTGGGATGCAGTGGCTGGAAGACACAAGGAAAATAAGCAA	1259
Db	1139	ACAGCGTGTGCTACTCTGTGGGAATGAGTGGCTGAAGTCTCAAGGAGACAAGCAGA	1198
2Y	1260	TGCTTTGCACGTGCTGGGCAACCGAGTCACTGTCGAAGAGACAGTGTAAACCCAGACTT	1319
Db	1199	TGCTGTGCATCTGCTGGCAATGGCGTCACTGTCAGGAGACAGCTGTGACCCAGACTT	1258
2Y	1320	ACGGTGGCAACTCAAAATGGAGGCAATGTGCTTACCAATTCACCTACAATGGCGAGACGT	1379
Db	1259	ACGGTGGCAACTCAAAACGGGAGCCCTGTGTCTTCCCGTTTCACTACACGGTAGACCT	1318
2Y	1380	TCTACTCTGCAACACAGAAGGGCGACAGGACGACATCTTGGTGCAGCAACACTTCGA	1439
Db	1319	TCCTACTCTGCACCAACCGAAGGGCGCAAGAGGACATCTGTGTGTAGCACAACCTTCAA	1378
2Y	1440	ATTTATGAGGAGCACAGAAATATCTCTTTTTCGCACAGACCACTATGTTTGGTTCAGATC	1499
Db	1379	ATTTATGAACAGACCAAGATTTCTTTCTCGACAGACCAACGGGGTTTTGGTTCAGATC	1438
2Y	1500	GAGGAGGAAATTCCAAATGGTGCCTTGTGCCACTTGTCCCTCTTCCCTATACAAACAACAAT	1559
Db	1439	GAGGTGGGAATTCCAAATGGTGCCTTGTGCCACTTCCCTCTTCTGTACAGCAACGGGAT	1498
2Y	1560	ACACTGATTCCTTTCAGGGCGAGAGACACAACATGAAGTGGTGGGACCAACAGAGA	1619
Db	1499	ACAGCACTGTACTTCTGAGGGTAGGGGGACAACATGAATGGTGGCGCACCAACCAGAGA	1558

QY	1620	ACTATGATGCCACACAGAAGTTTGGTTCCTGCCCCCATGGCTGCCCCACAGGAATCTGCA	1678
Db	1559	ACTACGATGCCATCAGAAGTTTGATCTTGCCCCAATGGCTGCCCATGAGGAGATCTGCA	1618
QY	1680	CAACAAATGAAGGGGTCACTATCCGCAATTGGAGATCAGTGGGATGAAGCAGCATGACATGG	1739
Db	1619	CGACCDAACGAAGGGTCACTATCGCATTTGGGACCCAGTTGGGATGAAGCAGCATGACCTGG	1678
QY	1740	GTCCACATGATGAGGTGCGACGTGTGTGTGGAAATGGTCTGGGGAATGGACATGCAATGGCT	1799
Db	1679	GCCACATGATGAGGTGCGACGTGTGTGTGGAAACGGCCGTGGCAATGGGGCTGCATCCCT	1738
QY	1800	ACTCGCAGCTTCGAGATCAGTGCATTTGTGTGATGACATCACTTACAATGTGAACGACACAT	1859
Db	1739	ACTCCAGCTTCGAGATCAGTGCATTCGTGTGATGACATTACTTACAACGTCAACGACACGT	1798
QY	1860	TCCACAAGCGTCATGAAGAGGGGCGACATGCTGAACTGTGATCATGCTTCGTTCGGTCAAGGTCCGGG	1919
Db	1799	TCCACAAGCGTCAACGAGGAGGACATATCTGAACTGTACTCTGCTTCGGTCAAGGCCGGG	1858
QY	1920	GCAGGTGGAAGTGTGATCCCGTCGACCAATGCCAGGATTCAGAGACTGGGACGTTTATC	1979
Db	1859	GCAGATGGAATGTGACCCCATCGACCGATGCCAGATTCAGAGACCCGCGACATTTTACC	1918
QY	1980	AAATTGGAGATTCATGGGAGAAGTATGTCCATGGTGTGAGATACCAAGTCTACTGCTATG	2039
Db	1919	AGATTGGTGACTCTCGGAGAAGTTTGTGCATGGTGTGATACCAAGTGTACTGTTAAG	1978
QY	2040	GCGTGGCATTTGGGAGTGGCATTTGCCAACCTTTACAGACCTTATCCAACTCAAGTGGTC	2099
Db	1979	GCGTGGCATTTGGGAGTGGCATTTGCCAGCTCTGCAGACCTTACCCAGGACAACTGGAC	2038
QY	2100	CTGTCCAAGTATTTATCACTGAGACTCCGAGTCAGCCCAACTCCCAACCCCATCCAGTGGGA	2159
Db	2039	CTGTTCAAGTAAATTATCACGAGACCCCCAGCCAGCCCAATTCCCAACCCCATCCAGTGGGA	2098
QY	2160	ATGCACACAGCCATCTCACATTTCCAAGTACATTTCTCAGGTGGAGACCTTAAAAATCTGT	2219
Db	2099	ATCCCCCGAGCCTTCACACATCACCAAGTACATTTCTCAGGTGGAGACCTTAAACCTCTA	2158
QY	2220	TAGCGCGTTTGGAAAGGAGCTACCATACCAAGGCCACTTAAACTCCTACACCATCAAAAGGCC	2279
Db	2159	CGAGTGCCTGGAAGGAGCTACCATTTCCAGGCCACCTTAACTCCTATACCATCAAGGCC	2218
QY	2280	TGAAGCCTGGTGTGTATACAGGGGCCAGTCTATAGCATTCAGCAGTAGTAGCGCCACCAAG	2339
Db	2219	TGACCCAGAGTGTGATCTACGAGGGACAGCTCATCAGCATTCAGCAGTAGTAGCGGCACCAAG	2278
QY	2340	AAGTGAAGCTGCTTTGACTTCACCAACACAGCACACGACACTGTGACACAGCAACACCG	2399
Db	2279	AAGTGAAGCTGCTTTGACTT---CACCAACAGCGCCAGCACCTGTGACCGCAACACAG	2335
QY	2400	TGACAGGAGAGACGACTCCCTTTTCTCCTCTCTGTGGCCACTTCTGAAATCTGTGACCCGAAA	2459
Db	2336	TGACTGGAGAGACTGCGCCCTTTTCTCCTGTTGTGGCCACTTCCGAATCTGTCACTGAAA	2395
QY	2460	TCCAGCGAGTACTGTTGTGTTCTCTGGGTCTCAGCTTCCGACACGCGTGTCCGATTC	2519
Db	2396	TCCAGCCAGCAGCTCTGTGGTCTCTCTGGGTCTCAGCTTCCGACACGCGTGTCCGATTC	2455
QY	2520	GGGTGGAATATGAGCTGAGTGAGAGGGAGATCAGCCACAGTACCTGGATCTTCCCAAGCA	2579
Db	2456	GAGTGGAGTACGAACATGACGAGAGGAGGAGATGAGCCTCAGTACCTTGATCTTCCCAAGCA	2515
QY	2580	CAGCCACTTCTGTGAACATCCCTGACCTGTTCTGGCCGGAATAATCAATTGTAAATGTCT	2639
Db	2516	CAGCCACTTCTGTGAACATTCCTGACCTGCTCCCGGGCAGAAAGTAGTACATCGTCAACGCTCT	2575
QY	2640	ATCAGATATCTGAGGATGGGACGAGTTTGATCTGCTACTTCCAAACAAACACAGCGC	2699
Db	2576	ATCAGATATCTGAAGAGGGGAAACAGAGCTTTGATCTCTGCTATCATCAGACATCAGCAC	2635
QY	2700	CTGATGCCCTCTGACCCGACTGTGGACCAAGTTGATGACACCTCAATTTGTTGCT	2759

b	2636	CTGATGCGCTCCAGACCCCTACTGTGGACCCAGGTTGATGACATTCATTTGTTGTCGAT	2695
y	2760	GGAGCAGACCCAGCGCTCCCATCACAGGGTACAGAAAGTCTATTGCGCATCAGTAGAAG	2819
b	2696	GGAGCAGACCCAGCGACCATATCACAGGGTACAGATTGTCTATTCACTTCAGTAGAAG	2755
y	2820	GTAGCAGCAGACACTCAACCTTCTGAACTGCAAACTCGGTACCCCTCAGTGACTTGC	2879
b	2756	GCAGTAGCAGAACTCAACCTTCTGAAACGGCCAACTCCGTCACTCCCTCAGCGACTGC	2815
y	2880	AAGCTGCTGTTCAGTATAACATCACTATCTATGCTGTGGAGAAATACAGAAAGTACAC	2939
b	2816	AGCCCGGTGTTCAGTACAACTCACTATCTATGCTGTGGAGGAACACAGAGAGCACAC	2875
y	2940	CTGTTGTCAATCAACAAGAAACCATGCGACCCACCGCTCAGATACAGTGCCTCTCCCA	2999
b	2876	CCGTTTTCATCCAGCAGGAGACTCTTGGCGTCCACGATCCGATGATGTTCCCGCTCCAA	2935
y	3000	GGACCTGCAGTTTGTGGAAGTGACAGACGTGAAGGTACCATCATGTGACACCCGCTG	3059
b	2936	AGGACCTACAGTTTGTGGAAGTGACCGACGTGAAAGTCACCATCATGTGACACCTCTTA	2995
y	3060	AGAGTGCAGTGACCGGCTACCGTGTGGATGTATCCCGTCAACCTCGCTGGCGAGCAAG	3119
b	2996	ATAGCGCAGTGACTGGATACCGTGTGGATGTCTCCCTGCTCAACTGCCAGGGGACATG	3055
y	3120	GGCAGGCTGCCCATCAGCAGAGAAACACTTGTGAGAAAGTCAACGGGCTGTCCCTGGGG	3179
ib	3056	GGCAGAGGCTGCCTGTCAACAGAGAAACACTTGTGTGAAGTCACCGGACTGTCCCAAGGG	3115
iy	3180	TCACCTATTACTTCAAAAGTCTTTGGCAGTGAGCATCCCGTCAACCTCGCTGGCGAGCA	3239
ib	3116	TCAGTACTCTTCAAAAGTCTTTGTGTGTCATCAGGCGAGGGGAAGCAAGCTTGACAG	3175
iy	3240	CTCAACAGACAAACAACTGGAATGCTCCCATTAACCTCCAGTTTGTCAATGAACATGATT	3299
ib	3176	CACAACAGACACCAAACTCGATGCTCCCACTTAACCTCCAGTTTGTCAATGAACACGACA	3235
iy	3300	CTACTGTCTCTGGTGAATGGACTCCACTCGCGGCCAGATAACAGGATACCGACTGACCG	3359
ib	3236	GAACAGTCTTGTTAACTTGGACTCCACTCGAGCCGGATAGCAGGCTACCGACTGACAG	3295
iy	3360	TGGGCTTTACCGAAGAGAGCAGCCAGAGCATGAANTGTGGTGCTCTGTCTCCAGT	3419
ib	3296	TGGGCTTCAGCCGAGGAGCGACCCCAAGCAGTCAATGTGGGACCATGGCTTTCAGT	3355
iy	3420	ACCCACTGAGGAATCTGCAGCTGCACTGAGTACACCGTATCCCTGTGGCCATAAAGG	3479
ib	3356	ATCCCTTGAGAAATCTGCAGCTGGGTCTGTAGTACACTGTGACCTTGATGGCTGTGAAG	3415
iy	3480	GCAACGAAGAGAGCCCAAGCCACTGGAGTCTTTACCACTGCGAGCTGGAGCTCTA	3539
ib	3416	GCAACCAAGCAGAGTCCCAAGCCAGCAGGAGTCTTTACTACCTGCGAGCTCTGGCTCCA	3475
iy	3540	TTCCACCTTCAACACCGAGGTGACTGAGACCAACCATTTGTGATCACATGAGCGGCTGCTC	3599
ib	3476	TTCCACCTTATACACCGAGGTGACAGAGACCAAACTGTGATCACCTGGAGCCCGCTC	3535
iy	3600	CAAGAAATGGTTTTAAGCTGGGTGTACGACCAAGCAGCGGAGGAGACCAACGAGAAG	3659
ib	3536	CAAGGAATGGCTTCAAGCTGGGTGTACGACCAAGCCAGGAGGTGAAGACCCCGAGAG	3595
iy	3660	TGACTTCAGACTCAGGAAGCATCGTTGTGTCCGGCTTGACTCCAGGAGTAGAATACGTCT	3719
ib	3596	TGACTTCAGACTCAGGAAGCATCGTTGTGTCTGGCTTGACTCCAGGCGTGAATACACGT	3655
iy	3720	ACACATCCAGTCTTGAGAGATGGACAGGAAGAGATGCCCAATGTGAACAAAGTGG	3779
ib	3656	ACACATCCAGTCTTGAGGAGCGGCAGAGAGAGATGCACAAATGTCAACCGAGTAG	3715
iy	3780	TGACACCAATTTGCTCCACCAACAAACTTGATCTGTGGAGCAAAACCTGCACATCGAGTGC	3839

3715	TGACACCGCTGCTCCCCCAACCAACTTGCACCTCGAGGCCCAATCCTGTACACATCTGGAGTGC	3775
3840	TCACAGTCTCTCTGGAGAGGAGCAGCACCCAGACATTTACTGGTTATAGAAATTTACACAA	3899
3776	TTACCGTCTCTCTGGAGAGGAGCACACCCCCAGATTTACTGGCTACAGAAATAACACACA	3835
3900	CCCTTACAAACGGCCAGCAGGAGGAAATTTCTTTGGAAGAAGTGGTCCATGCTGATCAGAGCT	3959
3836	CCCCACAAACGGCCAGCAGGAAACCGCTTTTGGAGAAGTGGTTTATGCGCATCAGAGTT	3895
3960	CCTGCACCTTTTGATTAACCTGAGTCCCGGCTGGAGTACAAATGTCTAGTGTTTACACTGTCA	4019
3896	CCTGCACTTTTGAACACCGTAATCTCTGGCTGGAGTACAAATGTCTAGTGTTTACACTGTCA	3955
4020	AGGATGACAAGAAAGTGTCCTTATCTCTGATACCATCATCCCA	4063
3956	AAGATGACAAGAAAGTGTCCTTATCTCTGATACCGTCAATCCAGAGGTGCCCCAGCTCA	4015
4064	-----	4063
4016	CTGACCTAAGCTTTGTTGTATATAAATGACTCAAGCATCGGCTGAGTGGACCCCGCTAA	4075
4064	-----	4063
4076	ACTTTCCACCAATTATCGGCTACCGAATCACAGTAGTTGGCGCAGGAGAAGGATCCCCA	4135
4064	-----	4063
4136	TTTTTGAAGATTTTGTGGACTCTCTAGTAGGATACTACAGTTACAGGCTGGAAACCG	4195
4064	-----	4063
4196	GCATTGACTATGACATCAGCGTTATCACTCTCATTTAATGGCGGAGAGAGTGCCCTACTA	4255
4064	-----GCTGTTCTCTCCCTCCCACTGACCTGGATTACCAACCAATTTGGTC	4106
4256	CATGACACAGCAACGGCGCTCCCTCTCCACAGATCTGCGATTCACCAATATCGTTC	4315
4107	CAGACACCATGCGTGTCACTGGGCTCCACCCCATCTCCATTTGATTTAAACCAATCTCTGG	4166
4316	CGACACATATGCGGGTCACTTTGGCGCCGCTCCGTCATTTAGCTAAACCAAGCTCTTGG	4375
4167	TGGGTTACTCACCTGTGAAAAATGAGAGATGTTGCAGAGTTGTCAATTTCTCTTCAG	4226
4376	TGGCTACTCACCTGTGAGAACGAGGAGGATGTGGCAGAGCTGTCCATTTACCCCTCAG	4435
4227	ACAATGAGTGGTCTTAAACAATTCCTGCTCGGTACAGAAATGTAGTGAAGTGTCTCCA	4286
4436	ACAACGGTGGTCTTAAACAATTCCTGCTCGGTACAGAAATGTAGTGAAGTGTCTCCA	4495
4287	GTGTTACGAAACAACATGAGACACACCTCTTAGAGGAAGACAGAAAAAGGCTTTGATT	4346
4496	CGGTGTACGAACAGCATGAGACATACCTCTCAGAGGAAGACAGAAAAAGGCTTGGACT	4555
4347	CCGCAACTGGCATTTTCTTCTGTATATTAATCGGCAACTTTTTTACTGTGCACTGGATTG	4406
4556	CCGCAACTGGTTTTGATTCTTCTGATGTCAACCCCACTCATTCACCGTCCACTGGGTGG	4615
4407	CTCCTCAGGCCACCATCACTCGGCTACAGGATCGCCATCATCCCGAGCACTTCAGTGGGA	4466
4616	CTCCTCGGGCCCCATCAACCGGTACATCATCCGCCATCAACCGCGAGCATTCCTCGCGAA	4675
4467	GACCTCGAGAAGATCGGGTGCACCACTCTCGGAATTTCCATCACCTTACCAACTCACTC	4526
4676	GACCCAGCAAGACCGAGTCCGCGCTTCAAGGAATTTCTATCACCCCTCACCAACTTAATC	4735
4527	CAGGCACAGAGTATGTGGTCAGCATCGTGTCTTAAATGGCAGAGAGGAAGTCCCTTAT	4586
4736	CGGCAACGGAGTACATTGTCCACATCATTTCTGTTAATGGCAGAGAGAGCCCCCAC	4795
4587	TGATTGCCCAACATCAACAGTTTCTTGATGTTCCGAGGACCTTGGAGTTGTTGCTCGGA	4646
4796	TGATTGCCCAAGCAATCCACGGTTTCGATGTCCGAGAGATCTGGAGGTCAATCGCTTTCCA	4855

2y	4647	CCCCACAGCTACTGATCAGCTGGGATGCTCTGTGTCAAGTGAGATATTTACAGGA	4706	5727	CTGTGTGACAGATGCTACTGAGACCAACCATCAACATTTAGCTGGAGAACCAAGACTGAGA	5786
Dd	4856	CCCCACAGCTCTGCTCATGTTGGGAACCCCGCCGCTCTGTGGCTATTACAGAA	4915	5936	CCCGTGTGACCGACGCTACAGAACTACCATCACTATTAGCTGGAGAACGAGACAGGA	5995
2y	4707	TCACCTACGAGAAACAGGAGGAATAGCCCTGTCAGGAGTTCACTGTGCCCTGGAGCA	4766	5787	CGATCACTGGCTCCAAAGTTGATCGGTTCCAGGCAATGGCCAGACTCCAATCCAGAGAA	5846
Dd	4916	TCACCTATGAGAGACAGGAGGAATAGCCCTGTCAGGAGTTCACTGTGCCCGAAGCA	4975	5996	CGATCACTGGCTTTCCAGTCTGATGCCATTCCAGCAATGGCCAGACCCCGGTTACAGGA	6055
2y	4767	AGTCTACAGCTACCATCAGCGGCTTAAACCTGGAGTTGATTACCATCACTGTGTATG	4826	5847	CCATCAAGCCAGATGTTCAGAAGCTACACCATCACTGGCTTACAAACAGGCACTGACTACA	5906
Dd	4976	AGTCCACGCCACCATCACAACATTAAACAGAGGAGACTACACCATCACCTGTATG	5035	6056	CCATCAGCCCGGATGTTCAGAAGCTATATACTATTACAGGTTTACAGCCAGCACTGACTACA	6115
2y	4827	CTGTCACTGGCGGTGGAGACAGCCGCCCAAGCAGCAAGCCAAATTCATTAAATACGAA	4886	5907	AGATCTACTGTACACCTTGAATGACAAATGCTCGAGCTCCCTGTGTGTCTATCAGCGCT	5966
Dd	5036	CTGTCACTGGCGGTGGAGACAGCTCCAGCAGCAGCAAGCCAGTTTCCATTCAATATCAAA	5095	6116	AGATCCACTGTACACGCTCAACGCAATGCCGAGCTCTCCGTGTCTATTGATGCT	6175
2y	4887	CAGAAATGACAAACCATCCAGATGCAAGTACCGGATGTTTCAGGACAAACAGATTAGTG	4946	5967	CCACTGCCATTGATCACCACCATCCAACTGCGTTTCTGGCCACCACACACCCCAATTCCTTGC	6026
Dd	5096	CAGAAATGACAAACCATCCAGATGCAAGTACCGGATGTTTCAGGACAAACAGATTAGTG	5155	6176	CCACGGCCATTGATGCCCATCCCACTCCAACTGCGTTTCTGACCAACACACCCCACTTGC	6235
2y	4947	TCAAGTGGCTGCTTCAAGTTCCCTGTACTGTGTTACAGAGTAAACCACTCCCAAAA	5006	6027	TGGTATCATGGCAGCGGCAAGTGCAGGATTACCGGTACATCATCAAGTATGAGAGC	6086
Dd	5156	TCAGTGGCTGCTTCAAGTTCCCTGTACTGTGTTACAGAGTAAACCACTCCCAAAA	5215	6236	TGGTATCATGGCAGCGCAGCGGTGCGGATTTACTGGCTTACATTATCAAGTATGAGAGC	6295
2y	5007	ATGACACGAGCAACAAACCTAAACTCAGTCCAGATCAACAGAAATGACTATTG	5066	6087	CTGGTCTCTCCACAGAGAAAGTGGTCCCTCGGCCCGCCCTGGTGTACAGAGGCTACTA	6146
Dd	5216	ATGGCTTAGGACCAACAAATCTCAAACTGTCACTCCAGATCAAAACAGAAATGACCATG	5275	6296	CTGGATCCCTCCACAGAGAGTGGTCCCTCGGCCCGCCCTGGTGTACAGGAGGCCACA	6355
2y	5067	AAGGCTTCAGCCACACAGTGAGTATGTGTTAGTGTCTATGCTCAGAAATCCAGCGGAG	5126	6147	TTACTGGCTCGAAACCGGAAACCGGAATATACATTTATGTCTATTCCTCGCTGAAGATATC	6206
Dd	5276	AAGGTTTTCACACCCAGCGTGAGTATGTGTTAGTGTCTATGCTCAGAAATCCAGCGGAG	5335	6356	TCACTGGTCTGGAGCCAGAACCGGATACACCATCTATGTATGTCATGCACTGAAGAAACAATC	6415
2y	5127	AGAGTACGCTCTGGTTCAGACTGAGTAAACCAATGATCGCCCTTAAGGACTGCGAT	5186	6207	AGAAGAGCAGGCGCTGATTGGAAGGAAAAAGACAGAGAGCTTCCCAACTGGTAAACC	6266
Dd	5336	AAAGCAGGCTGCTGGTTCAGACTGAGTAAACCAATGATCGCCCTTAAGGACTGCGAT	5395	6416	AGAAGAGTGAAGCCCTGATTGGAGGAAAAAGACAGATGAGCTTCCCACTGGTTAACC	6475
2y	5187	TCAGTGTGATGATGTCATTAATCCATCAAAATGTGTTGGAAAGCCACAGGGCAAGTTT	5246	6267	TTCCACACCCCAATCTTCATGACACAGAGATCTTGGATGTTCCTTCCACAGTTCAAAAGA	6326
Dd	5396	TCAGTGTGATGATGTCATTAATCCATCAAAATGTGTTGGAAAGCCACAGGGCAAGTTT	5455	6476	TTCCACACCCCAATCTTCATGACACAGAGATCTTGGATGTTCCTTCCCACTGGTCAAAAGA	6535
2y	5247	CCAGTACAGGCTGACCTACTTCAGGCGCTGAGGATGGAATCCATGAGTATTCCTGCGAC	5306	6327	CCCTTTTCGTGACCCACCGCTGGGTATGACACTGGAATGGTATTTCAGCTTCTGCGCACTT	6386
Dd	5456	CCAGTACAGGCTGACCTACTTCAGGCGCTGAGGATGGAATCCATGAGTATTCCTGCGCG	5515	6536	CCCTTTTCGTGACCCACCGCTGGGTATGACCCGGAATGGTATTTCAGCTTCTGCGCACT	6595
2y	5307	CTGATGTTGAAGAGACACTCGAGAGTGCAGGCTCAGACCGGTTCTGAGTACACAG	5366	6387	CTGCTCAGCAACCCAGTGTGGGCAACAAATGATCTTTGAGGAAATGTTTGGCGGA	6446
Dd	5516	CTGATGTTGAAGAGACACTCGAGAGTGCAGGCTCAGACCGGTTCTGAGTACACAG	5575	6596	CCACCAACAAACCCAGTGTGGGCAACAAATGATCTTTGAGGAAATGTTTGGCGGA	6655
2y	5367	TCAGTGTGTTGCCCTTGACAGATGATGGAGAGCCAGCCCTGATTGGAAACCAAGTCCA	5426	6447	CCACACCGCCCAACAGCGGCCACCCCAATAAGGCATAGGCCAAGACCATACCCGCGGAATG	6506
Dd	5576	TCAGTGTGTTGCCCTTGACAGATGATGGAGAGCCAGCCCTGATTGGAAATGTTTGGAAATG	5535	6656	CCACGCCACCCACTGCGGCCACCCCGCTCAGGCTTAGGCCAAGACCATACCTGCCGAATG	6715
2y	5427	CAGCTATTCTGCAACCACTGACCTGAAGTTCACTCAGGTACACCCACAGCCCTGAGCG	5486	6507	T-----	6507
Dd	5636	CAGCCATTCTCTGCGCAACCAATCTGAAGTTCACTCAGGTGTACCCACCACTTGTACTG	5695	6716	TAGATGAGGAGTGCATAATCGGTGATGTTCCAGGGAGACGTAGACTACCACTCTATC	6775
2y	5487	CCAGTGGACACCAACCAATGTTCAAGTCTCACTGGATATCGAGTGGGTGACCCCAAGG	5546	6508	-----AGGACAGAAAGCTCTCTCTCAGAGAA	6533
Dd	5696	CCAGTGGACACCGCCAGTGTAAAGTCTCACTGGCTACCGGTGCGGTGACCCCGAAGG	5755	6776	CTCATGTTCCGGGTCAATCCAAATGCCCTCTACAGGACAGAAAGCTCTCTCTCAGAGAA	6835
2y	5547	AGAAGACCGGACCAATGAAGAAATCAACCTTGCTCTGACAGCTCATCGGTGTTGTAT	5606	6534	CCATCTCATGGGCCCCCATTTCCAGGACACTTTGAGTACATCAATTTTCATGCTCATCTGTTG	6593
Dd	5756	AGAAGACCGGACCAATGAAGAAATCAACCTTTCTCCAGACAGACCTCCGTGATTGTG	5815	6836	CCATCTCTTGGACGCCATTTCCAGGAGAGTTCTGAGTACATCATTTTCATGCCAACCTGTTG	6895
2y	5607	CAGGACTTATGGTGCCACCAATATGAAGTGAAGTGTCTATGCTCTTAAGGACACTTTGA	5666	6594	GCACTGATGAAGAACCTTTACAGTTTCCAGGTTTCTGGAACTTCTACAGTGCACCTCTGA	6653
Dd	5816	CAGGCTCATGTTGCGCCACCAAGTATGAAGTGAAGTGTCTATGCTCTCAAGGACACTTTGA	5875	6896	GCACTGACGAAGAGCCCTTACAGTTTCCAAAGTTTCCGGAAGCTTCTACAGTGCACCTCTGA	6955
2y	5667	CAAGCAGACCACTCAGGAGTTGTCAACCTCTGGAGAAATGTACGCCCAACAGAGAGGG	5726	6654	CAGGCTCTCAGGAGTGCACCTCAACATCATAGTGGAGGCACTGAAAGACCAAGCAGA	6713
Dd	5876	CAAGCAGACCACTCAGGAGTGTGTCAACCTCTGGAGAAATGTACGCCCTTCCAGAGAGGG	5935	6956	CTGGCTTACAGAGAGGGGTCACTTACAAATCATATAGTGGAGGCCCTGCACACAGAGGA	7015
2y				6714	GGCATAGGTTTCGGGAAGAGGTTGTTTACCGTGGGCAACTCTCTGTCAAGAGAGGTTGAACC	6773

1b	7016	GGCACAAGGTC	CCGAGAAGAGGTTGTTACTGTAGGCAACACTGTCAACCAAGGCGCTGAACC	7075
1y	6774	AACCTACGGATGACTCGTCTTGTGACCCCTACACAGTGTCCCATATTATGCGCGTGTGGAGATG	6833	
1b	7076	AGCCTACGGATGACTCATGCTTTGACCCCTTACACCGTTCCTCCATTAACCGCTGTGGAGAGG	7135	
1y	6934	AGTGGGAACGAATGTCTGTAATACAGGCTTTAAACTCTGTTGTGCCAGTCTTTAGCTTTGGAA	6993	
1b	7136	AATGGGACGGTTTATCTGATCTCTGGCTTTAAGCTCACTTGGCCAGTGTCTGGGCTTTGGCA	7195	
1y	6994	GTGGTCAATTTTCAGATGTGATTCACTCFAGATGTGTGCCATGACAATGTGTGAACTACAGA	6953	
1b	7196	GTGGTCAATTTTCAGATGCGGATTCACTTAATGTTGCCATGACACGTTGTCACTACAGA	7255	
1y	6954	TTGGAGAGAAAGTGGGACCGTCAAGGAGAAAAATTTGGCCAGATGATGAGCTGCMCATGTCTTG	7013	
1b	7256	TCGGAGAGAAAGTGGGATTCGTCAGGAGAAAAATTTGGCCAGCGGATGAGCTGCAATGTCTCG	7315	
1y	7014	GGACCGAAAGAGAGAAATTCAGATGTGACCCCTCATGAGCAACGTTTACGATGATGGGA	7073	
1b	7316	GGAAATGGAAAGGAGAAATTCAAATGCGATCCCATGTAAGCAACGTTTATGACGACGGGA	7375	
1y	7074	AGACATACCACTAGGAGAAACAGTGGCAGAAAGAAATATCTCGGTGCCAATTTGCTCTCGCA	7133	
1b	7376	AGACCTACCACTAGGAGAAACAGTGGCAGAAAGAGTATCTCGGAGCCATTTGCTCTCGCA	7435	
1y	7134	CATGCTTTGGAGCGCAGCGGGGCTGCGCTGTGACAACTGGCGAGACCTGGG--CGTG	7190	
1b	7436	CGTGTTCGGGGGCGAGCGGGGCTGCGCTGTGACAACTGGCGAGACCTGGGGCTGCTG	7495	
1y	7191	AACCCAGTCCCGAAGGCACTACTGGCCAGTCTCTCAACACAGTATTTCTCAGAGATACCATC	7250	
1b	7496	AACCCAGTCCCGATGGTATCCACTGGGCCACACCTCAACACAGTATACACAGAGATACCATC	7555	
1y	7251	AGAGAAACAAACATAATGTTAAATTTGCCCAATTTAGTGTCTTCATGCTTTAGATCTACAGG	7310	
1b	7556	AGAGAAACAAACATAATGTTAAATTTGCCCAATTTGAATGCTTTATGCCGTGTGACGTGAGG	7615	
1y	7311	CTGACAGAGAAGATTTCCGAGAGATAAATCATCTTTTCAATCCAGAGGAACCAAGCATCTCT	7370	
1b	7616	CTGACAGAGATGATTTCCAGAGAGTAAATCTTTTCCATCCAGGCCCAAGCCAAACAG--TGCTCT	7673	
1y	7371	CTCTGCCAAGATCCATCTAACTGAGTGAATGTTAGCAGAC---CCAGCTTAGAGTTCCTT	7427	
1b	7674	CTCTACCAAGGTCAATCCACACCCCACTGATGTAGCAGACCCCTCAATTTCTGAGTGGCT	7733	
1y	7428	CTTCTCTTTCTTAAGCCCTTTGCTCTCGAGGAAGTTCCTCCAGCTTCAGCTCAACTCACAGC	7487	
1b	7734	ATTTTCACTTTAAGCTTTCTGCTCTGGAGTCAAGTTCCTCAGCTTCAGCTCAACTTACAGC	7793	
1y	7488	TTCTCCAAAGCATCAACCTG--GGAGTTTCTGAGGGTTTCTCATAAATGAGGGCTGCACA	7546	
1b	7794	TTCTCCAAAGCATCCCGCCGGGATGTTTGAACATTCCTCTTAAATGGTGAGCAGTTGG	7853	
1y	7547	TGTCCTGTTCTGCTTCGAAGTATTCAAATACCGCTCAGTATTTTAAATCAAGTGAATCTAA	7606	
1b	7854	TGCCCTGTTCTGCTTTACGGGTATTCAGTACTGCTCAGTATTAATGCTCAAGAGAATCAAA	7913	
1y	7607	-----GATTTGGTTTGGGATTCAAATAGGAAGCATATGCGGCCAACCAAGATGCAAA	7657	
1b	7914	AGTTCTTGATTTTGGTCTGGGATCAA--AGGGAAACACAGGTAGCCAAACCAAGTGC--AA	7971	
1y	7658	TGTTTGTGAATGATATGACCAAAATTTTAAATAGGAAGTCAACCCAAAACATCTTCGCTTT	7717	
1b	7972	TGAATTGAATGCTAGTACCACAGCGGGAGCAGGAAGTTAAACACAGACAGTCTTCGCTTT	8031	
2y	7718	CACTT	7722	
2b	8032	CTTTT	8036	

```

US-10-191-803-113
; Sequence 113, Application US/10191803
; Publication No. US20040014040A1
; GENERAL INFORMATION:
; APPLICANT: MENDRICK, Donna
; APPLICANT: PORTER, Mark
; APPLICANT: JOHNSON, Kory
; APPLICANT: HIGGS, Brandon
; APPLICANT: CASTLE, Arthur
; APPLICANT: ELASHOFF, Michael
; TITLE OF INVENTION: Cardiotoxin Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5090US
; CURRENT APPLICATION NUMBER: US/10/191,803
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US 60/303,819
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/305,623
; PRIOR FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: US 60/369,351
; PRIOR FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: US 60/377,611
; PRIOR FILING DATE: 2002-05-06
; NUMBER OF SEQ ID NOS: 1140
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 113
; LENGTH: 8329
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20040014040A1 NM_019143
US-10-191-803-113

```

Query Match	68.3%; Score 5320.6; DB 15; Length 8329;
Best Local Similarity	81.8%; Pred. No. 0;
Matches 6484; Conservative	0; Mismatches 1044; Indels 397; Gaps 14;
184	GGATTTTTCGCGAGGTTTGGGAAGGACAAACTTGGTGCAACTTGCCTCCCGGTGCG 243
Db	123 GGACTTTTCTGAGGCTGGAGGGGAA-CCAACTTCGTGCCACTAGCTCTCTGGAG 181
Qy	244 GGGCTCTCTCC-CCACCGTCTCAACAATGTTAGGGTTCGGGGCCCG- --GGTGTGTGC 299
Db	182 GGGACTCTCTCCATCCACTCAAGATGCTCAGGGGTCCGGACCGCGGTGTGTGC 243
Qy	300 TGTTCGGCTCTGTGCTCGGTGGGACACGGGTGCCTCCACGGGAGCTCGAAGACGAGA 359
Db	242 TGCTAGCAGTCTGTGTGCTGGGACATCGTGGCTGCACGGAACCGGGAAGACGAAGA 301
Qy	360 GGCAAGGCTCAGCAAAATGGTTTCAGCCCCAGTCCCGGTGGTGTGTCAGTCAAGACGAAGCCCG 419
Db	302 GGCAGGCTCAGCAANTCGTGCAGCTCCGTCCCGGTGGTGTGTCAGTCAGAGCAAGCCTG 361
Qy	420 GTTGTATTGACAATGGAAAAACATATCAGATAAATCAACAGTGGGAGCGGACCTACCTAG 479
Db	362 GCTGTTTGTGACACGGGAAGCATTTATCAGATAAATCAGACGTGGGAACGGACCTACCTAG 421
Qy	480 GCAATGCGTGTGGTTGTACTGTTATCGAGGAACCGCAGGTTTAACTGCAGAAAGTAAC 539
Db	422 GCAACGGCCCTGGTTTGTACCTGCTATGGAGGAACAGAGGTTTAACTGCCGAGACAAAGC 481
Qy	540 CTGAAGCTGAAGAGACTTGCTTTGACAAGTACACTGGGAACACTTACCGAGTGGGTGACA 599
Db	482 CTGAACCTGAAGAGACCTGTTTTCACAAATACACTGGAAACACTTACAAAGTGGGTGACA 541
Qy	600 CTTATGAGCGTCTAAAGACTCCCATGATCTGGGACTGTACCTGCATCGGGGCTGGCCAG 659
Db	542 CTTATGAGCGCCTTAAAGATTCCCATGATCTGGGACTGTACCTGCATTTGGGGCTGGCCAG 601
Qy	660 GGAGAAATAAGCTGTACCATCGCAACCGCTGCCATGAAGGGGTCAGTCCCTACAAGATTG 719
Db	602 GCAGGATCAGCTGTGACCAATTCCTGCCAATGAAGGGGGTCAGTCCCTACAAGATTG 661
Qy	720 GTGACACCTGGAGGAGACCAATCAGAGACTGGTGGTTTACATGTTAGAGTGTGTGTCTTG 779

Db 662 GTGCAAGTGGAGAGGCCACATGAGACTGGTGGCTATATGTTGGAGTGTGTTGCTGG 721
Qy 780 GTAATGGAAGAGAGAAATGGACCTGCAAGCCATAGCTGAGAAAGTGTGTTGATCATCTG 839
Db 722 GGAATGGAAAGGGGAATGGACCTGCAAGCCATAGCTGAGAAATGTTTGGATCAGCTG 781
Qy 840 CTGGGACTTCTCTATGTTGGTGGAGAAAGCTGGAGAGACCCCTACCAAGGCTGGATGATG 899
Db 782 CTGGGACTTCTCTATGTTGGTGGAGAAAGCTGGAGAGACCCCTACCAAGGCTGGATGATG 841
Qy 900 TAGATTGTACTTGGCTGGGAG 959
Db 842 TGGACTGTACTTGGCTGGGAG 901
Qy 960 GCAAGATCAGGACACAGGACATCTCTATAGAAATGGAGAGACCTGGAGAGAGAGAGATA 1019
Db 902 GCAATGATCAGGACACAGGAG 961
Qy 1020 ATCGAGGAAACCTGCTCCAGTGCATCTGCACAGGCAAGCGGCGAGGAGAGTGGAGTGTG 1079
Db 962 ACAGAGGAAACCTGCTCCAGTGCATCTGCACAGGCAAGCGGCGAGGAGAGTGGAGTGTG 1021
Qy 1080 AGAGCACACCTCTGTGTGAGACACATCGAGGAGATCTGGCCCTTCCAGATGTTGCTG 1139
Db 1022 AGCGACA--TGTTCTACAGAGTGTCTTCCAGTGCATCTGGCTCTTCCAGATGTTGCGAA 1078
Qy 1140 CAGCTGTTTACCAACCGCAGCTCAACCGGAGCTCTCTCCCTATGGCCACTGTGTACAG 1199
Db 1079 CAGCTATTTTACCAACCGGAG 1138
Qy 1200 ACAGTGTGTGTGTACTCTGTGGGATGCAATGAGTGGTGAAGACACAAAGGAAATAGCAAA 1259
Db 1139 ACAGCGTGTGTGTACTCTGTGGGATGCAATGAGTGGTGAAGTCTCAAGGAGACAGCAGA 1198
Qy 1260 TGCCTTGCAGCTGCTTGGGACAGAGTCACTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1319
Db 1199 TGCTGTGCACTTGGCTGGGCAATGGCGTCACTGCGGAGAGAGAGAGAGAGAGAGAGAGAG 1258
Qy 1320 ACGGTGGCAACTCAATGGAGAGCAGATGTCTTACCAATTCACCTTACCAATGGCAGGACGT 1379
Db 1259 ACGGTGGCAACTCAATGGAGAGCAGATGTCTTACCAATTCACCTTACCAATGGCAGGACGT 1318
Qy 1380 TCTACTCTGACACACAG 1439
Db 1319 TCTACTCTGACACACAG 1378
Qy 1440 ATTATGAGCAGGACAGAGAAATACTCTTTCTGCACAGACCACTGTTTGGTTTTCAGACTC 1499
Db 1379 ATTATGAGCAGGACAGAGAAATACTCTTTCTGCACAGACCACTGTTTGGTTTTCAGACTC 1438
Qy 1500 GAGGAGGAAATTCGAATGGTGTGCTGCGCACTTCCCTTCTCTATACAAACACCAATTT 1559
Db 1439 GAGGTGGGAAATTCGAATGGTGTGCTGCGCACTTCCCTTCTCTATACAAACACCAATTT 1498
Qy 1560 ACACGTGTTGCACTTCTGAGGAG 1619
Db 1499 ACAGGCACTGTACTTCTGAGGAG 1558
Qy 1620 ACTATGATCCGACAGAGAGTGTGGTGTGCTGCGCACTTCCCTTCTCTATACAAACACCAATTT 1679
Db 1559 ACTAGATGCCGATCAGAGAGTGTGGTGTGCTGCGCACTTCCCTTCTCTATACAAACACCAATTT 1618
Qy 1680 CAAACCAATGAGGGGTCTATGACCGCATTTGGAGATCAGTGGGATGAGCAGATCAGATGG 1739
Db 1619 CGAACCAAGAGGGGTCTATGATCCATTTGGGACCACTGGGATGAGCAGATCAGATGG 1678
Qy 1740 GTACATGATGAGGTGCAAGTGTGTGGGAAATGGTGTGGGAAATGGACATGCAATTCCT 1799
Db 1679 GCAATGATGAGGTGCAAGTGTGTGGGAAATGGTGTGGGAAATGGACATGCAATTCCT 1738
Qy 1800 ACTCGAGCTTGCAGATCAGT 1859

Db 1739 ACTCCAGCTCCGAGATCAGTGTGATCGTGTGATGACATTTACTTACAAAGTCAACGACAGT 1798
Qy 1860 TCCAAAGGCTCATGAAAGAGGGGCACTGCTGAATGTGTAATGTCTTCCGTTCAGGGTCCGG 1919
Db 1799 TCCAAAGGCTCATGAAAGAGGGGCACTGCTGAATGTGTAATGTCTTCCGTTCAGGGTCCGG 1858
Qy 1920 GCAAGTGGAGTGTGATCCGTCGACCAATGCGAGGATTCAGAGACTGGGAGCTGTTTATC 1979
Db 1859 GAGATGGAAATGTGACCCATCGACCGATGCCAAGATTCAGAGACCCGACATTTTACC 1918
Qy 1980 AAATGGAGATTCATGGGAGAAATGTGCTGATGTGTGATGATGATGATGATGATGATGATG 2039
Db 1919 AGATTGGTCACTCTCGGAGAAATTTGTGCATGTGTGATGATGATGATGATGATGATGATG 1978
Qy 2040 GCGTGGCATTCGGGAGTGGCATTCGCAACCTTTACAGACCTTATCAAGCTTCAAGTGTGTC 2099
Db 1979 GCGTGGCATTCGGGAGTGGCATTCGCAACCTTCTGCAAGCTTACCCAGGACCACTGGAC 2038
Qy 2100 CTGTGGAAGTATTTATCACTGAGATCTCGAGTCAAGCCAACTCCCAACCCCATCCAGTGA 2159
Db 2039 CTGTTCAAGTAAATTCACGGAG 2098
Qy 2160 ATGCACCAAGCATCTCAATTTCCAAAGTACATTTCTCAGGTGGAGACCTAAAAATCTG 2219
Db 2099 ATGCCCCGAGGCTTTCACACATCAACAGTACATTTCTCAGGTGGAGACCTAAAACTCTA 2158
Qy 2220 TAGGCGTGTGGAGGAGAGTACCATACCAAGGCACTTAACTCTCTACACCATCAAGGCT 2279
Db 2159 CGGTCGCTGGAGGAGAGTACCATTCAGGAGCACTTAACTCTCTATACCATCAAGGCT 2218
Qy 2280 TGNAGCTGTGTGTATACGAGGGGCGGCTCATCAGATCCAGAGTACGAGGAGAGAGAGAGAG 2339
Db 2219 TGACCCGAGGCTGTGATCTACGAGGAGAGAGCTCATCAGCATCCAGAGTACGAGGAGAGAG 2278
Qy 2340 AAGTGAAGTGTGATCTTCACTTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2399
Db 2279 AAGTGAAGTGTGATCTTCACTTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2335
Qy 2400 TGACAGGAGAGAGAGTCCCTTCT 2459
Db 2336 TGACTGGAGAGAGTGGGCGCTTCT 2395
Qy 2460 TCACAGCAGTGTGTGTGTCT 2519
Db 2396 TCACAGCAGAGAGTGTGTGTCT 2455
Qy 2520 GGTGGAATATGATGTGATGAGGAG 2579
Db 2456 GAGTGGAGTACGAATGAGGAG 2515
Qy 2580 CAGCAGTCTGTGAGACATCCCTGAGCT 2639
Db 2516 CAGCAGTCTGTGAGACATCCCTGAGCT 2575
Qy 2640 ATCAGATATCTGAGGAG 2699
Db 2576 ATCAGATATCTGAGGAG 2635
Qy 2700 CTGATGCCCTCTGAG 2759
Db 2636 CTGATGCCCTCTGAG 2695
Qy 2760 GGAGCAG 2819
Db 2696 GGAGCAG 2755
Qy 2820 GTAGCAGCAG 2879
Db 2756 GCAATGAGCAG 2815
Qy 2880 AACCTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2939
Db 2816 AGCCCGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2875

2940 CTGTGTTCATTCAACAGAGAAACCACTGGGACCCCACTCGCTCAGATACAGTGCCTCTCCCA 2999
 b CCGTTTTCATCCAGCAGGAGACTCTGGGTCCTCCAGATCCGATGATGTCCCGTCCAA 2935
 Y GGGACCTCAGTTTGTGGAAGTGACAGAGCTGAAGTCAACATCATGTGGACACCGCTG 3059
 b AGACCTACAGTTTGTGGAAGTGACCGAGCTGAAGTCAACATCATGTGGACACCTCCTA 2995
 Y AGAGTCACTGACCGGCTACCGTGTGATGTGATCCCGTCAACCTGCTGGCGAGCAG 3119
 b ATAGCGACGTGACTGATACCGTGTGATGTGATCCCGTCAACCTGCTGGCGAGCAG 3055
 Y GGCAGAGGCTGCTGTCAACAGGAACCTTTGCTGAGTCAACCGAGTGTCCCGAGGG 3115
 b TCACCTATTACTTCAAGTCTTTGAGTGAAGCACTGGGAGGAGAGCAAGCTCTGACTG 3239
 b TCAGGTACCTCTTCAAGTCTTTGCTGTGATCAGGGCAGGGAAGCAAGCTCTGACAG 3175
 Y CTAACAGACACCAACTGATGCTCCCACTAACTCCAGTTTGTCAATGAATCTGATT 3299
 b CAAACAGACCAACTGATGCTCCCACTAACTCCAGTTTGTCAATGAATCTGATT 3235
 Y CTAAGTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3359
 b GAACAGTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3295
 Y TGGGCTTACCGGAGAGGCAAGCCAGAGTACAAATGTGGGACCAATGTGGGTTCCAA 3419
 b TGGGCTTACCGGAGAGGCAAGCCAGAGTACAAATGTGGGACCAATGTGGGTTCCAA 3355
 Y ACCCACTGAGGATCTGACGCTGATCTGATGATGATGATGATGATGATGATGATGATG 3479
 b ATCCCTGAGAACTCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3415
 Y GCAACAGAGAGCCCAAGCCACTGAGTCTTTTACCACTGAGGCTGAGGCTGCTGCTG 3539
 b GCAACAGAGAGTCTCAAGCCCAAGCCACTGAGTCTTTTACCACTGAGGCTGAGGCTG 3475
 Y TTCACCTTACAAACCGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATG 3599
 b TTCACCTTACAAACCGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATG 3535
 Y CAAGAAATGGTTTAACTGGGTGATGATGATGATGATGATGATGATGATGATGATGATG 3659
 b CAAGAAATGGTTTAACTGGGTGATGATGATGATGATGATGATGATGATGATGATGATG 3595
 Y TGACTTCAGACTCAGGAAGCATCGTTGTGCTGGCTTGACTCCAGGATGATGATGATGAT 3719
 b TGACTTCAGACTCAGGAAGCATCGTTGTGCTGGCTTGACTCCAGGATGATGATGATGAT 3655
 Y ACACCAATCCAAAGTCTGAGGAGTGAAGAGAGATGAGGAGATGAGGAGATGAGGAGAT 3779
 b ACACCAATCCAAAGTCTGAGGAGTGAAGAGAGATGAGGAGATGAGGAGATGAGGAGAT 3715
 Y TGACACCATGCTCTCCACCAACAACTTGATCTGGAGGCAACCTGACTGAGGAGTGC 3839
 b TGACACCATGCTCTCCACCAACAACTTGATCTGGAGGCAACCTGACTGAGGAGTGC 3775
 Y TCACAGTCTCTGGGAGAGAGACCAACCCAGCATTTACTGGTTATAGAAATACCAAA 3899
 b TTACCGTCTCTGGGAGAGAGACCAACCCAGCATTTACTGGCTACAGAAATACCAAA 3835
 Y CCGCTACAAAGCCGACAGGAAATCTTTGAGAGAGTGGTCCATGCTGATGAGCT 3959
 b CCGCTACAAAGCCGACAGGAAATCTTTGAGAGAGTGGTCCATGCTGATGAGCT 3895
 Y CCGTGCATTTTGAATCACTGAGTCCCGGCTGGAGTCAATGTGATGTTTACCTGTC 4019
 b CCGTGCATTTTGAATCACTGAGTCCCGGCTGGAGTCAATGTGATGTTTACCTGTC 3995

QY 4020 AGGATGACAGGAAGTGTCCCTATCTCTGTATACCATCATCCCA 4063
 Db 3956 AGATGACAGGAAGTGTCCCTATCTCTGTATACCATCATCCAGAGTGTCCCGCTCA 4015
 QY 4064 4063
 Db 4016 CTGACCTAAGCTTTTGTATATAACTGACTCAAGCATCGGCTGAGTGGAGCCCGCTAA 4075
 QY 4064 4063
 Db 4076 ACTCTCCACCATATTCGGGTACCGAATCACAGTAGTTGCGGCAGAGAGGGATCCCCA 4135
 QY 4064 4063
 Db 4136 TTTTGAAGATTTTGTGACTCCTCAGTAGGATACTACAGATTACAGGCTGGAACCCG 4195
 QY 4064 4063
 Db 4196 GCATGTGATGACATCAGCGTTATCCTCTCATTAATGCGGAGAGAGTGCCCTACTA 4255
 QY 4064 4063
 Db 4256 CACTGACACAGCAACCGGCGTCCCTCTCCACGGATCTCGATTTCACCAATATCGGTC 4315
 QY 4107 CAGACACATCGGTGTCACTGGGTCCACCCCATCCATTTGATTTAAACCACTCTCTG 4166
 Db 4316 CGGACATATGCGGGTCACTTGGGCGGCGCTCCGTCATTTGAGCTAACCACTCTG 4375
 QY 4167 TGCCTTACTCCTCTGTAATAATGAGGAAGATTTGACAGATTGTCAATTTCTCTTCA 4226
 Db 4376 TGCCTTACTCCTCTGTAATAATGAGGAAGATTTGACAGATTGTCAATTTCTCTTCA 4435
 QY 4227 ACATGAGTGTCTTAAACAAATCTCTCTGCTGTACAGATATGATGAGTGTCTTCA 4286
 Db 4436 ACACGCGGTGTCTTAAACAAATCTCTCTGCTGTACAGATATGATGAGTGTCTTCA 4495
 QY 4287 GTGTCTACGAACCAATGAGGACACCTCTTACAGGAAGACAGAAAAACAGGCTCTTGA 4346
 Db 4496 GCGTGTACGAACCAATGAGGACACCTCTTACAGGAAGACAGAAAAACAGGCTCTG 4555
 QY 4347 CCGCACTGAGTGTCTTAAACAAATCTCTCTGCTGTACAGATATGATGAGTGTCTTCA 4406
 Db 4556 CCGCACTGAGTGTCTTAAACAAATCTCTCTGCTGTACAGATATGATGAGTGTCTTCA 4615
 QY 4407 CTCTCTGAGCAACCATCACTGCTGTACAGGATCGCCATCATCCGAGCACTTCACTGGGA 4466
 Db 4616 CTCTCTGAGCAACCATCACTGCTGTACAGGATCGCCATCATCCGAGCACTTCACTGG 4675
 QY 4467 GACTCGAGAGATTCGGGTGCGGCTCTCGAAATTCCTTCACTCACTCACTCACTCTC 4526
 Db 4676 GACTCGAGAGATTCGGGTGCGGCTCTCGAAATTCCTTCACTCACTCACTCACTCTC 4735
 QY 4527 CAGGACAGAGATTCGGGTGCGGCTCTCGAAATTCCTTCACTCACTCACTCACTCTC 4586
 Db 4736 CCGGACAGAGATTCGGGTGCGGCTCTCGAAATTCCTTCACTCACTCACTCACTCTC 4795
 QY 4587 TGAATGCGCAACATCAACAGTTCCTGATGTTCCGAGGACCTGGAAAGTGTGCTCGGA 4646
 Db 4796 TGAATGCGCAACATCAACAGTTCCTGATGTTCCGAGGACCTGGAAAGTGTGCTCGGA 4855
 QY 4647 CCGGACAGAGATTCGGGTGCGGCTCTCGAAATTCCTTCACTCACTCACTCACTCTC 4706
 Db 4856 CCGGACAGAGATTCGGGTGCGGCTCTCGAAATTCCTTCACTCACTCACTCACTCTC 4915
 QY 4707 TCACCTTACGAGAAACAGGAGGAAATAGCCCTGTCTCAGGAGTTCATCTGCTGGGAGCA 4766
 Db 4916 TCACCTTACGAGAAACAGGAGGAAATAGCCCTGTCTCAGGAGTTCATCTGCTGGGAGCA 4975
 QY 4767 AGTCTACAGCTACCATCAGGCTTAAACCTGAGTTGATTTACCATCACTGCTGATG 4826
 Db 4976 AGTCTACAGCTACCATCAGGCTTAAACCTGAGTTGATTTACCATCACTGCTGATG 5035
 QY 4827 CTGTCTGCGGCTGGAGACAGAGCCCGCCGCAAGCAAGCAATTTCCATTAATTTACCGAA 4886

Db 5036 CTGTCACTGGCGTGGGACAGTCCAGCCAGCAGCAAGCCAGTTTCCATCATTAATAA 5095
 Qy 4887 CAGAAATTGACAAAACATCCAGATGCAAGTACCGATGTTTCAGGACAAAGCATAGTG 4946
 Db 5096 CAGAAATTGACAAAGCAATCCAGATGCAAGTACCGATGTTTCAGGACAAAGCATAGTG 5155
 Qy 4947 TCAAGTGGCTGCTTTCAAGTTCCTCTGTTACTGTTTACAGAGTAAACCACTCCAAAA 5006
 Db 5156 TCAGGTGGCTGCTTTCACTCTCTGTTGACAGGTTACAGAGTGAACACCGCTCCAAAA 5215
 Qy 5007 ATGGACAGGACCAAAAAAATCAAACTGACAGTTCAGATCAAAAGAAATGATATTG 5066
 Db 5216 ATGGCTAGGACCAAAAAATCAAACTGACAGTTCAGATCAAAAGAAATGATATTG 5275
 Qy 5067 AAGGCTTGACGCCACAGTGGAGTATGTTAGTGTCTATGCTCAGAAATCAAGCGGAG 5126
 Db 5276 AAGGTTTGCAACCCACCGTGGAGTATGTTAGTGTCTATGCTCAGAAACCGGACGGAG 5335
 Qy 5127 AGAGTCAGGCTCTGTTGACATGACATGACGATCAACAACTTATGCTGACGCTTAAAGACTGGCAT 5186
 Db 5336 AAAGCCAGCCCTGGTTTACAGTTCAGATGACGATGACCAACATTTGACCGCCCTTAAAGACTGGCAT 5395
 Qy 5187 TCACATGATGTGATGTCCATTCATCAAAATTTGCTTGGAAAGCCACAGGGGCAAGTTT 5246
 Db 5396 TCACATGATGTGATGTCCATTCATCAAAATTTGCTTGGAAAGCCACAGGGGCAAGTTT 5455
 Qy 5247 CCAGTACAGGCTGACCTACTCGAGCCCTGAGGATGGAATCCATGAGTATTTCCCTGAC 5306
 Db 5456 CCAGTACAGGCTGACCTACTCGAGCCCTGAGGATGGAATCCATGAGTATTTCCCTGAC 5515
 Qy 5307 CTGATGGTGAGAACACATGACAGTGCAGAGGCTCAGACCGGTTCTGAGTACACAG 5366
 Db 5516 CTGATGGTGAGAACACATGACAGTGCAGAGGCTCAGACCGGTTCTGAGTACACAG 5575
 Qy 5367 TCAGTGTGGTTGCTTGCATGACATGATGGAGCCAGCCCTGATTTGGAACCCAGTCCA 5426
 Db 5576 TCAGTGTGGTTGCTTGCATGACATGATGGAGCCAGCCCTGATTTGGAATGCTCAGTCCA 5635
 Qy 5427 CAGCTATTTCTGACCAACTGACCTGAGTTCACTCAGGTACACCCACAGCCCTGAGCG 5486
 Db 5636 CAGCCATTTCTGCGCAACCAATCTGAAAGTTCACTCAGGTGTCCACCCACCTTGAAGT 5695
 Qy 5487 CCCAGTGCACACCACTGTTTCACTGAGTATCGAGTGGGGTGACCCCAAGG 5546
 Db 5696 CCCAGTGCACACCGCTGTTTCACTGAGTATCGAGTGGGGTGACCCCAAGG 5755
 Qy 5547 AGAAGCCGACCAATGAAAGAAATCAACCTTGTCTTGCAGCTCATCCGTGGTTGAT 5606
 Db 5756 AGAAGACAGGACCAATGAAAGAAATCAACCTTGTCTTGCAGACAGCACTCCGTGATGTTGT 5815
 Qy 5607 CAGGACTTATGGTGGCCACCAATATGAAGTGAAGTGTCTATGCTTTAAGGACACTTTGA 5666
 Db 5816 CAGGCTCATGGTGGCCACCAAGTATGAAGTCAAGCTCTATGCTTCAAGGACACTTTGA 5875
 Qy 5667 CAGGACAGCAGCTCAGGAGTTGTACCACTCTGGAGAAATGTACGCCCAACAGAGGG 5726
 Db 5876 CAGGACAGCAGCTCAGGAGTTGTACCACTCTGGAGAAATGTACGCCCAACAGAGGG 5935
 Qy 5727 CTCGTGTGACAGATGCTACTGAGACCAATCAACCATTTAGCTGGAGAACCAAGACTGAGA 5786
 Db 5936 CCGGTGTGACGACGCTACAGAACTCAATCACTATTTAGCTGGAGAACCAAGACTGAGA 5995
 Qy 5787 CGATCACTGGCTTCAAGTTGATGCGTTCCAGCCAAATGGCCAGACTCCAAATCCAGAGAA 5846
 Db 5996 CGATCACTGGCTTCAAGTTGATGCGTTCCAGCCAAATGGCCAGACTCCAAATCCAGAGAA 6055
 Qy 5847 CCATCAAGCCAGATGTCAGAGCTACCATCAGCTTACACCCAGGCACTGACTACA 5906
 Db 6056 CCATCAAGCCGATGTCAGAGCTTACTATTAAGGTTTACGCCAGGCACTGACTACA 6115
 Qy 5907 AGATCTACCTGTACCTTGAATGATGCTCGGAGCTCCCTGTGCTATCGACGCT 5966

6116 AGATCCACCTGTACACGCTCAACGACAAATGCCGGAGCTCTCCTGTGGTCAITGATGCC 6175
 Qy 5967 CCACTGCCATTGATGACACCATCCAACTGCTGTTCTGSCCACCACACCAATTCCTTGC 6026
 Db 6176 CCAAGCCATTGATGCCCATCCAACTGCTGTTCTGACCAACACCACTCTCTTGC 6235
 Qy 6027 TGTATCATGCGACGCGCCACGTCGCCAGGATTACCGGCTTACATCAATAGTATGAAAGC 6086
 Db 6236 TGTATCATGCGACGCGCCACGTCGCCAGGATTACCGGCTTACATCAATAGTATGAAAGC 6295
 Qy 6087 CTGGGTCTCTCCACAGAGTGGTCCCTGCGCCCGCCCTGCTGTCACAGAGGCTACTA 6146
 Db 6296 CTGGATCCCTCCACAGAGTGGTCCCTGCGCCCGCCCTGCTGTCACAGAGGCTACTA 6355
 Qy 6147 TTAAGTGGCTTGGAAACCGGAAACCGAATATACAAATTTATGCTATTTGCTGAAAGATAATC 6206
 Db 6356 TCAGTGGTCTGGAGCCAGGAAACCGAGTACACCATCTATGTCATCGCACTGAAGAAATC 6415
 Qy 6207 AGAAGACGAGCCCTGATTTGGAAGAAAAAGACAGACAGCTTCCCACTGGTAACCC 6266
 Db 6416 AGAAGATGAGCCCTGATTTGGAAGAAAAAGACAGATGAGCTTCCCACTGGTTACCC 6475
 Qy 6267 TTCCACACCCCAATCTTCATGACACAGATCTTGGATGTTCTTCCACAGTTCAAAAGA 6326
 Db 6476 TTCCACACCCCAATCTTCATGACACAGATCTTGGATGTTCTTCCACAGTTCAAAAGA 6535
 Qy 6327 CCGCTTTCGTACCCACCTCGGTATGACACTGGAATGGTATTCAGCTTCTCGGCACTT 6386
 Db 6536 CCGCTTTCGTACCCACCTCGGTATGACACTGGAATGGTATTCAGCTTCTCGGCACTT 6595
 Qy 6387 CTGGTCAGCAACCCAGTGTGGGCAACAAATGATCTTTCAGGAAACATGTTTTCAGCGGA 6446
 Db 6596 CCACCAACCAACCCAGTGTGGGCAACAAATGATCTTTCAGGAAACATGTTTTCAGCGGA 6555
 Qy 6447 CCACCGCCCAACCAACCCAGTGTGGGCAACAAATGATCTTTCAGGAAACATGTTTTCAGCGGA 6506
 Db 6656 CCACCGCCCAACCAACCCAGTGTGGGCAACAAATGATCTTTCAGGAAACATGTTTTCAGCGGA 6715
 Qy 6507 T----- 6507
 Db 6716 TAGATGAGGAGTCCAAATCGGTCATGTTCCCGGGAGACGTAGACTACCACTCTATC 6775
 Qy 6508 -----AGGCAAGAAAGCTCTCTCTCAGACAA 6533
 Db 6776 CTCATGTTCCGGGCTCAATCCAAATGCTCTACAGGACAAAGAGCTCTCTCTCAGACAA 6835
 Qy 6534 CCACTCATGGGCCCCATTCCAGGACACTCTGAGTACATCATTTCACTGTCATCTGTTG 6593
 Db 6836 CCACTCTTGGAGCCCAATTCAGGAGAGTCTGAGTACATCATTTCACTGTCATCTGTTG 6895
 Qy 6594 GCATGATGAAGAACCCCTTACAGTTCAGGTTCCGAACTTCTACCACTGCTGCACTCTGA 6653
 Db 6896 GCATGACGAGAGCCCTTACAGTTCGAACTTCTGGAACCTTCTACCACTGCTGCACTCTGA 6955
 Qy 6654 CAGGCTCACCAGAGTGCCTTACAAATCATATGAGGAGGACCTGAAGACCAAGCAGAGA 6713
 Db 6956 CTGGCTTACAGAGGGTCACTTACAAATCATATGAGGAGGCTGCAACCAAGAGAGA 7015
 Qy 6714 GGCATAAGGTTCCGGAAGAGGTTGTCCTGGGCAACTCTGTCACAGAAAGGCTTGAAC 6773
 Db 7016 GGCATAAGGTTCCGGAAGAGGTTGTCCTGGGCAACTCTGTCACAGAAAGGCTTGAAC 7075
 Qy 6774 AACCTACGAGTACTGCTGTTGACCCCTTACAGTGTCCCAATTTATGCGTGTGAGATG 6833
 Db 7076 AGCTACGAGTACTGCTGTTGACCCCTTACAGGTTTCCCAATTTACGCGTGTGAGAGG 7135
 Qy 6834 AGTGGAAAGAAATGCTGAAATCAGGCTTTTAAATGTTGTCGAGTGTCTTTCAGCTTGGAA 6893
 Db 7136 AATGGAGCGGTTATCTGACTCTGGCTTTTAACTCACTTGGCCAGTGTCTTGGCTTGGCA 7195
 Qy 6894 GTGGTCACTTACAGATGATCATCTAGATGTTGTCATGACAAATGTTGTAACACTACAAGA 6953
 Db 7196 GTGGTCACTTACAGATGATCATCTAAATGTTGTCATGACAAACGTTGTAACACTACAAGA 7255

Y 6954 TTGGAGAGAGTGGACCGTCAAGGAGAAATGGCCAGATGATGAGCTGCACATGTCTTG 7013
b 7256 TCGGAGAGAGTGGATCGTCAAGGAGAAATGGCCAGATGAGCTGCACATGTCTCG 7315
Y 7014 GGAACGGAAAGGAGAAATCAAGTGTGACCTCATGAGGCAACGTGTACGATGATGGGA 7073
b 7316 GGAATGGAAGGGAGAAATCAAAATCGCATCCCATGAAGCAACGTGTATGACGACGGGA 7375
Y 7074 AGACATACACATGAGAGACAGTGGCAGAGGATATCTGGTGCCATTTGGCTCCTGCA 7133
b 7376 AGACCTACACATGAGAGACAGTGGCAGAGGATATCTGGAGCCATTTGGCTCCTGCA 7435
Y 7134 CATGCTTTGGAGGCCAGCGGGCTGGCGTGTGACAACCTGCCGACAGACTGG---GGTG 7190
b 7436 CGTGTTCGGGGGCCAGCGGGCTGGCGTGTGACAACCTGCCGACAGACTGGGGCTGCTG 7495
Y 7191 AACCCAGTCCGAGGACATGAGTGGCCAGTCTGACCAACAGTATTTCTCAGAGATACCATC 7250
b 7496 AACCCAGTCCGAGGATGATCCAGAGATATCTTCCATCCAGCCCAAGCCCAACAG--TGCT 7555
Y 7251 AGAGAACAAACACTAATGTTAATTGCCCAATTTGAGTGTCTTCAATGCTTTAGATGTACAGG 7310
b 7556 AGAGAACGACACTAATGTTAATTGCCCAATTTGAGTGTCTTCAATGCTTTGAGCTGTGAGG 7615
Y 7311 CTGACAGAGAGATCCCGAGAGTAAATCATCTTTCCATCCAGAGGACAGCATGTCT 7370
b 7616 CTGACAGAGATGATCCAGAGATATCTTCCATCCAGCCCAAGCCCAACAG--TGCT 7673
Y 7371 CTCTGCCAAGATCCATCTAACTGGAGTGTATGTAGCAGAC---CCAGCTTAGAGTTCTT 7427
b 7674 CTCTACCAAGTCAATCCACACCCAGTGTATGTAGCAGACCTCCATTTCTGAGTGTGTC 7733
Y 7428 CTCTCTTTCTTAAGCCCTTGTCTCGAGAGAGTCTCCAGCTTCAGCTCAACTCAGAC 7487
b 7734 ATTTCAACCTTAAGCCTTCTGCTCTGAGTCAAGTTCTCAGCTTCAGCTCAACTTACAGC 7793
Y 7488 TTCTCCAAGCATCACCGTG-GGAGTTTCTGAGGGTTTCTCATAAATGAGGGCTGCACA 7546
b 7794 TTCTCCAAGCATCGCCCGGGATGTTTGGACTTCCCTCTTAAATGGTGACAGTTGG 7853
Y 7547 TTGCTGTCTTCTGATCGAAGTATTCATACCGCTCAGTATTTTAAATGAAGTATCTTAA 7606
b 7854 TGCCCTGTCTGCTTCAGGGTATTCAGTACTGCTCAGTATTTATTTCTAAGAGATCAAA 7913
Y 7607 -----GATTTGGTTGGGATCAATAGGAAGCATATGAGCCCAACCAAGATGCAAA 7657
b 7914 AGTCTGTGATTTGGTCTGGATCAA-AGGGAACACAGGTAGCCACACGATGC-AA 7971
Y 7658 TGTTTGAAATGATGACCAAAATTTTAAAGTAAAGTACCCAAACACTTCTGCTTT 7717
b 7972 TGAATTGAATGGTAGTACCCCAAGAGCGGGAGCAGGAAGTTAAACCAGACAGTTCTGCTTT 8031
Y 7718 CACTT 7722
b 8032 CTTT 8036

Search completed: March 8, 2004, 04:36:35
Job time : 2579 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

DM nucleic - nucleic search, using sw model

Run on: March 10, 2004, 09:19:42 ; Search time 69.2294 seconds
(without alignments)
12152.440 Million cell updates/sec

Title: US-10-084-817-8

Perfect score: 1516

Sequence: 1 cttgtttttggacatagct.....gctgaaataaaatgncgtc 1516

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA.*

- 1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	1515	99.9	1516	4	US-09-919-172-38
2	1387.8	91.5	1399	4	US-09-023-655-1342
3	332.8	22.0	1001	4	US-09-641-638-86
4	319.4	21.1	1362	4	US-09-023-655-1152
5	268.4	17.7	1339	4	US-09-023-655-1149
6	250.2	16.5	2305	3	US-08-526-136-1
7	250.2	16.5	2311	3	US-08-526-136-3
8	245.2	16.2	1301	4	US-09-023-655-1190
9	239.8	15.8	957	1	US-08-125-746-4
10	239.8	15.8	1460	6	5225537-3
11	239.8	15.8	1567	1	US-08-125-746-2
12	236.6	15.6	981	4	US-09-324-096A-1
13	236.6	15.6	981	4	US-09-324-096A-3
14	236.6	15.6	981	4	US-09-324-096A-5
15	232.8	15.4	2137	4	US-09-976-594-348
16	216	14.2	7938	3	US-09-331-581-14
17	215.2	14.2	7286	3	US-09-331-581-3
18	184.6	12.4	1356	4	US-09-010-147B-15
19	184.6	12.4	1356	4	US-09-641-638-80
20	158	10.4	1793	4	US-09-620-312D-747
21	153.2	10.1	644	3	US-09-328-111-721
22	141.2	9.3	502	3	US-09-328-111-625
23	141.2	9.3	630	3	US-09-328-111-491
24	135.4	8.9	501	3	US-09-328-111-481
25	125.4	8.3	1293	4	US-09-325-932A-17
26	124.4	8.2	1001	4	US-09-641-638-106
27	124.4	8.2	1001	4	US-09-641-638-107

C	28	118.8	7.8	643	3	US-09-385-982-306	Sequence 306, App
	29	116.6	7.7	741	4	US-09-641-638-81	Sequence 81, Appl
	30	102.2	6.7	543	3	US-09-328-111-428	Sequence 428, App
	31	101.2	6.7	616	4	US-09-621-976-1120	Sequence 1120, Ap
	32	101.2	6.7	813	4	US-09-641-638-100	Sequence 100, App
	33	101.2	6.7	943	4	US-09-641-638-99	Sequence 99, Appl
	34	100.6	6.6	396	4	US-09-640-173-60	Sequence 60, Appl
	35	100.6	6.6	396	4	US-09-713-550-60	Sequence 60, Appl
	36	99.2	6.5	789	4	US-09-325-932A-20	Sequence 20, Appl
	37	95.4	6.3	903	4	US-09-641-638-83	Sequence 83, Appl
	38	83.2	5.5	1001	4	US-09-641-638-115	Sequence 115, App
	39	82.4	5.4	1409	4	US-09-976-594-349	Sequence 349, App
	40	72.2	4.8	442	4	US-09-621-976-16600	Sequence 16600, A
	41	70	4.6	401	4	US-09-643-597-261	Sequence 261, App
	42	70	4.6	401	4	US-09-480-884A-261	Sequence 261, App
	43	70	4.6	401	4	US-09-542-615A-261	Sequence 261, App
	44	70	4.6	401	4	US-09-606-421B-261	Sequence 261, App
	45	61.4	4.1	1000	4	US-09-641-638-85	Sequence 85, Appl

ALIGNMENTS

RESULT 1

US-09-919-172-38
; Sequence 38, Application US/09919172
; Patent No. 6673545
; GENERAL INFORMATION:
; APPLICANT: Farris, Mary
; APPLICANT: Turner, Christopher M.
; TITLE OF INVENTION: PROSTATE CANCER MARKERS
; FILE REFERENCE: PA-0036 US
; CURRENT APPLICATION NUMBER: US/09/919,172
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/222,469
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PERL Program
; SEQ ID NO 38
; LENGTH: 1516
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673545 1303785CB1
; NAME/KEY: unsure
; LOCATION: 1512
; OTHER INFORMATION: a, t, c, g, or other
US-09-919-172-38

Query Match 99.9%; Score 1515; DB 4; Length 1516;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	CTTTGTTTGGACATAGCTGAGCCATGTACTTCAACAGAGGCGAGCCAACTTACTACT 60	
Db	1	CTTTGTTTGGACATAGCTGAGCCATGTACTTCAACAGAGGCGAGCCAACTTACTACT 60	
Qy	61	CTCGTTGTTAGTGTGGCTTCTTTTAAATCTATATAATCAGAGCCCAAGTCTCCAC 120	
Db	61	CTCGTTGTTAGTGTGGCTTCTTTTAAATCTATATAATCAGAGCCCAAGTCTCCAC 120	
Qy	121	TGCCAGTGTGAATCTTCAGAGAGAAATCTCTTTAGTCTTTTTCGCAAGAGGTAGAT 180	
Db	121	TGCCAGTGTGAATCTTCAGAGAGAAATCTCTTTAGTCTTTTTCGCAAGAGGTAGAT 180	
Qy	181	AAAGACACTTTTCAAAAAATGCAATGGTATCAGAAATCTCTCAAGCAGCCCTGGTTTATT 240	
Db	181	AAAGACACTTTTCAAAAAATGCAATGGTATCAGAAATCTCTCAAGCAGCCCTGGTTTATT 240	
Qy	241	GAATAATGAAGAGCAGGAATATGTTCAACTGTGAAGTCACTCAAAAGTGTGTCGCCGATCA 300	
Db	241	GAATAATGAAGAGCAGGAATATGTTCAACTGTGAAGTCACTCAAAAGTGTGTCGCCGATCA 300	

QY	301	GGGGTGAGGCCCTATCCTTACCTTCCATCCTCGGATGTCGTGCCTTGCATAAGGCC	360
DB	301	GGGGTGAGGCCCTATCCTTACCTTCCATCCTCGGATGTCGTGCCTTGCATAAGGCC	360
QY	361	ATAATGGTTTAAAGGTTGGATGAGTGAAGCAACCATCATTTCTAATCTCTAATGAGGAAACAAAT	420
DB	361	ATAATGGTTTAAAGGTTGGATGAGTGAAGCAACCATCATTTCTAATCTCTAATGAGGAAACAAAT	420
QY	421	GCACAGCGTCAACAGATCAAAAGCAGCATATCTCCAGGAAACAGGAAAGCCCTCGATGAA	480
DB	421	GCACAGCGTCAACAGATCAAAAGCAGCATATCTCCAGGAAACAGGAAAGCCCTCGATGAA	480
QY	481	ACACTGAAGAAAGCCCTTACAGGTCACTTGAAGGTTGTTTATAGTCTGCTATAAACT	540
DB	481	ACACTGAAGAAAGCCCTTACAGGTCACTTGAAGGTTGTTTATAGTCTGCTATAAACT	540
QY	541	CCAGCGCAATTTGATGCTGATGAACCTTCGTGTCATGAAGGCGCTTGGAACTGATGAA	600
DB	541	CCAGCGCAATTTGATGCTGATGAACCTTCGTGTCATGAAGGCGCTTGGAACTGATGAA	600
QY	601	GATACTCTAATTCAGATTTTGGCATCAAGACTCAAAAGAAATCAGAGACATTAACAGG	660
DB	601	GATACTCTAATTCAGATTTTGGCATCAAGACTCAAAAGAAATCAGAGACATTAACAGG	660
QY	661	GTCTACAGAGAGAACTGAAGAGAGATCTGGCCAAAGACATAACCTCAGACATCTGGA	720
DB	661	GTCTACAGAGAGAACTGAAGAGAGATCTGGCCAAAGACATAACCTCAGACATCTGGA	720
QY	721	GATTTTCGGAAACGTTTGCTTCTTCTGCTAAGGTTGACCGATCTGAGACATTCGTGTG	780
DB	721	GATTTTCGGAAACGTTTGCTTCTTCTGCTAAGGTTGACCGATCTGAGACATTCGTGTG	780
QY	781	AATGAAGACTTGCTGATTCAGATGCGCAGGCGCTTGATGAAGCAGGAGAAAGGAAAG	840
DB	781	AATGAAGACTTGCTGATTCAGATGCGCAGGCGCTTGATGAAGCAGGAGAAAGGAAAG	840
QY	841	GGACACAGCTAAACGTTGTTCAATACCATCTTACCACAGAGCTATCCAACTTCGC	900
DB	841	GGACACAGCTAAACGTTGTTCAATACCATCTTACCACAGAGCTATCCAACTTCGC	900
QY	901	AGAGTGTTCAGAAATACACCAAAGTACAGTAAGCATGACATGAACAAAGTTCTGACCTG	960
DB	901	AGAGTGTTCAGAAATACACCAAAGTACAGTAAGCATGACATGAACAAAGTTCTGACCTG	960
QY	961	GAGTTGAAAGGTGACATGTGAATGCGCTCACAGCTATCGTGAAGTGGCCCAAGCAAA	1020
DB	961	GAGTTGAAAGGTGACATGTGAATGCGCTCACAGCTATCGTGAAGTGGCCCAAGCAAA	1020
QY	1021	CCAGCTTTCTTTGCAGAGAAGCTTCATCAAGCCATGAAAGGTTTGGAACTCGCCATAAG	1080
DB	1021	CCAGCTTTCTTTGCAGAGAAGCTTCATCAAGCCATGAAAGGTTTGGAACTCGCCATAAG	1080
QY	1081	GCATTGATCAGGATTTATGGTTTCCGTTCTGAAATTTGACATGATGATCAAGCATTC	1140
DB	1081	GCATTGATCAGGATTTATGGTTTCCGTTCTGAAATTTGACATGATGATCAAGCATTC	1140
QY	1141	TATCAGAAGATGTATGTTATCTCCCTTTGCCAAGCCATCCTGGATGAAACCAAAGGAGAG	1200
DB	1141	TATCAGAAGATGTATGTTATCTCCCTTTGCCAAGCCATCCTGGATGAAACCAAAGGAGAG	1200
QY	1201	TATGAGAAATCCCTGGTGGCTTTTGTGGAGGAAACTTAAACATTCCTTGGATGCTCAA	1260
DB	1201	TATGAGAAATCCCTGGTGGCTTTTGTGGAGGAAACTTAAACATTCCTTGGATGCTCAA	1260
QY	1261	GCTATGATCAGAAGACTTTAATTATATATTTTCACTCTATAAGCTTAAATAGGAAAGTTT	1320
DB	1261	GCTATGATCAGAAGACTTTAATTATATATTTTCACTCTATAAGCTTAAATAGGAAAGTTT	1320
QY	1321	CTTCAACGAGTATCAGTGTAGCTACCTACATGCTGTAAGAAATATAGCCTTTAATCATTTT	1380
DB	1321	CTTCAACGAGTATCAGTGTAGCTACCTACATGCTGTAAGAAATATAGCCTTTAATCATTTT	1380

QY	1381	TTATATTATAACTCTGTATATAAGAGATAAGTCCATTCTTTTAAATAATGTTTTCCCCCAAAAC	1440
DB	1381	TTATATTATAACTCTGTATATAAGAGATAAGTCCATTCTTTTAAATAATGTTTTCCCCCAAAAC	1440
QY	1441	CATAAAACCTTATACAAGTGTGTTCTAGTAGTAAACAATACATGAGAAAGATGCTATGTAGCTG	1500
DB	1441	CATAAAACCTTATACAAGTGTGTTCTAGTAGTAAACAATACATGAGAAAGATGCTATGTAGCTG	1500
QY	1501	AAAAATAAAATGNGCTC	1516
DB	1501	AAAAATAAAATGNGCTC	1516

RESULT 2
 US-09-023-655-1342
 ; Sequence 1342, Application US/09023655
 ; Patent No. 6607879
 ; GENERAL INFORMATION:
 ; APPLICANT: Cocks, Benjamin G.
 ; APPLICANT: Susan G. Stuart
 ; APPLICANT: Jeffrey J. Seilhamer
 ; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
 ; TITLE OF INVENTION: EXPRESSION
 ; NUMBER OF SEQUENCES: 1508
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
 ; STREET: 3174 PORTER DRIVE
 ; CITY: PALO ALTO
 ; STATE: CALIFORNIA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/023,655
 ; FILING DATE: HEREWITH
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Zeller, Karen J.
 ; REGISTRATION NUMBER: 37,071
 ; REFERENCE/DOCKET NUMBER: PA-0001 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (650) 855-0555
 ; TELEFAX: (650) 845-4166
 ; INFORMATION FOR SEQ ID NO: 1342:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1399 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: GENBANK
 ; CLONE: g34387
 ; US-09-023-655-1342

	Query Match	91.5%	Score 1387.8	DB 4	Length 1393
	Best Local Similarity	99.8%	Pred. No. 0	Indels 3	Gaps 0
	Matches 1389	Conservative 0	Mismatches 0		
QY	125	AGTCGAAATCTTCAGAGAAGATTTCTCTTTAGTCTCTTTTCACAGAGAGGTAGAGATAAG	184		
DB	1	AGTCGAAATCTTCAGAGAAGATTTCTCTTTAGTCTCTTTTCACAGAGAGGTAGAGATAAG	60		
QY	185	ACACTTTTCAAAAATGGCAATGTTATCAGAAATTCCTCAAGCAGGCCTGGTTTATTTGAAA	244		
DB	61	ACACTTTTCAAAAATGGCAATGTTATCAGAAATTCCTCAAGCAGGCCTGGTTTATTTGAAA	120		

2Y	245	ATGAAGAGCAGGAATATGTTCAAACTGTGAAGTCATCCAAAAGTGCGTCCGGATCAGCGG	304
2B	121	ATGAAGAGCAGGAATATGTTCAAACTGTGAAGTCATCCAAAAGTGCGTCCGGATCAGCGG	180
2Y	305	TGAGCCCCATCCTACCTTCAATCATCCTCGGATGTCGCTTGCCTTCATTAAGGCCATAA	364
2B	181	TGAGCCCCATCCTACCTTCAATCATCCTCGGATGTCGCTTGCCTTCATTAAGGCCATAA	240
2Y	365	TGGTTAAAGGTTGGATGAAGCAACCATCATTTGACATTCTAACTTAAGCGAACCAATCCAC	424
2B	241	TGGTTAAAGGTTGGATGAAGCAACCATCATTTGACATTCTAACTTAAGCGAACCAATCCAC	300
2Y	425	ACGGTCAACAGATCAAGCAGCATATCTCCAGGAAACAGGAGCCCTCGATGAACAC	484
2B	301	ACGGTCAACAGATCAAGCAGCATATCTCCAGGAAACAGGAGCCCTCGATGAACAC	360
2Y	485	TGAAGAAACCCCTTACAGCTCACCTTGAAGGAGTTGTTTTAGCTCTGCTAAAACTCCAG	544
2B	361	TGAAGAAACCCCTTACAGCTCACCTTGAAGGAGTTGTTTTAGCTCTGCTAAAACTCCAG	420
2Y	545	CAGCAATTTGATCTGATGAACCTTCGTGTCGATGAAGGCTTGGACTGATCAAGATA	604
2B	421	CAGCAATTTGATCTGATGAACCTTCGTGTCGATGAAGGCTTGGACTGATCAAGATA	480
2Y	605	CTCTAATTCAGATTTTGGCATCAAGAATTAACAAGAAATCAGAGACATTAACAGGGTCT	664
2B	481	CTCTAATTCAGATTTTGGCATCAAGAATTAACAAGAAATCAGAGACATTAACAGGGTCT	540
2Y	665	ACAGAGAGAACTGAAGAGAGATCTGGCCAAAGACATAACCTCAGACACATCTCGAGATT	724
2B	541	ACAGAGAGAACTGAAGAGAGATCTGGCCAAAGACATAACCTCAGACACATCTCGAGATT	600
2Y	725	TTTCGAAACCCCTTGGCTTCTCTTGTAAAGGTTGACCGATCTGAGGACTTTTGGTGTGAATG	784
2B	601	TTTCGAAACCCCTTGGCTTCTCTTGTAAAGGTTGACCGATCTGAGGACTTTTGGTGTGAATG	660
2Y	785	AGACTTTGCTGATTCAGATGCGAGGCTTGTATGAAGCAGGAGAAAGGAGAAAGGGA	844
2B	661	AGACTTTGCTGATTCAGATGCGAGGCTTGTATGAAGCAGGAGAAAGGAGAAAGGGA	720
2Y	845	CAGACGTAAACGTTTCAATACCATCCCTTACCACAGAGAGCTATCCACAACTTCGCAGAG	904
2B	721	CAGACGTAAACGTTTCAATACCATCCCTTACCACAGAGAGCTATCCACAACTTCGCAGAG	780
2Y	905	TGTTTCAGAAATACACCAAGTACAGTAGCATGACATGAACAAAGTTCTGACCTGGAGT	964
2B	781	TGTTTCAGAAATACACCAAGTACAGTAGCATGACATGAACAAAGTTCTGACCTGGAGT	840
2Y	965	TGAAGGTCACATTGAGAATGCTCACAGCTATCGTGAAGTGCGCCACAAAGCAACAC	1024
2B	841	TGAAGGTCACATTGAGAATGCTCACAGCTATCGTGAAGTGCGCCACAAAGCAACAC	900
2Y	1025	CTTTCTTTGAGAGAGCTTCATCAAGCCATGAAGGTTTGGAACTCGCCATAGGCAT	1084
2B	901	CTTTCTTTGAGAGAGCTTCATCAAGCCATGAAGGTTTGGAACTCGCCATAGGCAT	960
2Y	1085	TGATCAGGATTATGGTTTCCCGTCTGTAATTCATGAATGATCAAGATCAAGCTTCTATC	1144
2B	961	TGATCAGGATTATGGTTTCCCGTCTGTAATTCATGAATGATCAAGATCAAGCTTCTATC	1020
2Y	1145	AGAAGATGATGATGATCTCCCTTTGCAAGCCATCTCGATGAACCAAGAGAGATG	1204
2B	1021	AGAAGATGATGATGATCTCCCTTTGCAAGCCATCTCGATGAACCAAGAGAGATG	1080
2Y	1205	AGAAATCTGGTGGCTCTTGTGGAGGAACTTAACATTCCTCTGATGTTCTCAAGCTA	1264
2B	1081	AGAAATCTGGTGGCTCTTGTGGAGGAACTTAACATTCCTCTGATGTTCTCAAGCTA	1140
2Y	1265	TGATCAGAGACTTTAATATATATTTTCACTTAAGCTTAAATAGAAAGTTTCTTC	1324
2B	1141	TGATCAGAGACTTTAATATATATTTTCACTTAAGCTTAAATAGAAAGTTTCTTC	1200

QY	1325	AACAGGATTACAGTGTAGCTACCTACCTACCTGCTGAAAAATATAGCCTTTAAATCATTTTTAT	1380
DB	1201	AACAGGATTACAGTGTAGCTACCTACCTACCTGCTGAAAAATATAGCCTTTAAATCATTTTTAT	1260
QY	1385	ATTATAAATCTGTGTAATAGAGATAAGTCCATTTTTTAAAAAATGTTTTTCCCACCAACCATTA	1444
DB	1261	ATTATAAATCTGTGTAATAGAGATAAGTCCATTTTTTAAAAAATGTTTTTCCCACCAACCATTA	1320
QY	1445	AAACCCCTATACAAGTTGTTCTTAGTAAACAATACATGAGAAAGATGCTCTATGTAGCTGAAAA	1504
DB	1321	AAACCCCTATACAAGTTGTTCTTAGTAAACAATACATGAGAAAGATGCTCTATGTAGCTGAAAA	1380
QY	1505	TAAAAATGNCCTC 1516	
DB	1381	TAAAAATGNCCTC 1392	

RESULT 3
US-09-641-638-86
; Sequence 86, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Amick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; FILE REFERENCE: GENSET.051CP1
; CURRENT APPLICATION NUMBER: US/09/641,638
; CURRENT FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 86
; LENGTH: 1001
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 501
; OTHER INFORMATION: 10-245-412 : polymorphic base A or G
; NAME/KEY: misc.binding
; LOCATION: 481..500
; OTHER INFORMATION: 10-245-412.mis1, potential
; NAME/KEY: misc.binding
; LOCATION: 502..521
; OTHER INFORMATION: 10-245-412.mis2, potential complement
; NAME/KEY: primer.bnd
; LOCATION: 90..107
; OTHER INFORMATION: upstream amplification primer
; NAME/KEY: primer.bnd
; LOCATION: 509..528
; OTHER INFORMATION: downstream amplification primer, complement
; NAME/KEY: misc.binding
; LOCATION: 489..513
; OTHER INFORMATION: 10-245-412 potential probe
; NAME/KEY: misc.feature
; LOCATION: 544,852,855..857,865
; OTHER INFORMATION: n=a, g, c or t
US-09-641-638-86

Query March 22, 0%; Score 332.8; DB 4; Length 1001;
Best Local Similarity 99.1%; Pred. No. 7,2e-92;
Matches 334; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1180 CTGGATGAACCAAGGAGAGTATGAGAAATCTCTGGTGGCTCTTTGTGAGGAACTAA 1239
Db 21 CAGGATGAACCAAGGAGATTTGAGAAATCTCTGGTGGCTCTTTGTGAGGAACTAA 80
QY 1240 ACATTCCTCTGATGCTCAAGCTATGATCAGAGACTTTAAATATATATTTTCACTTA 1299
Db 81 ACATTCCTCTGATGCTCAAGCTATGATCAGAGACTTTAAATATATATTTTCACTTA 140
QY 1300 TAAGCTTAAATAGGAAAGTTTCTTCAACAGGATTACAGTGTAGCTACATGCTGAAA 1359
Db 141 TAAGCTTAAATAGGAAAGTTTCTTCAACAGGATTACAGTGTAGCTACATGCTGAAA 200
QY 1360 AATATAGCTTTAAATCAATTTTATATATATATATATATATATATATATATATATAT 1419
Db 201 AATATAGCTTTAAATCAATTTTATATATATATATATATATATATATATATATATAT 260
QY 1420 TTAATAATGTTTCCCAACCAATTAACCAATTAACCAATTAACCAATTAACCAATTA 1479
Db 261 TTAATAATGTTTCCCAACCAATTAACCAATTAACCAATTAACCAATTAACCAATTA 320
QY 1480 AGAAAGATGCTATGATAGCTGAAATTAATAATGNCGTC 1516
Db 321 AGAAAGATGCTATGATAGCTGAAATTAATAATGNCGTC 357

RESULT 4

US-09-023-655-1152
; Sequence 1152, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Sellhauer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; NUMBER OF INVENTIONS: 1
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1152:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1362 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g187137
; US-09-023-655-1152

Query Match 21.1%; Score 319.4; DB 4; Length 1362;
Best Local Similarity 59.0%; Pred. No. 1.1e-87;
Matches 567; Conservative 0; Mismatches 391; Indels 3; Gaps 1;
QY 302 CGGTGAGCCCTATCTACCTTCAATTCATCTCGATGTGCTCCCTTGATAGGCA 361
Db 126 CTGTCAAAAGCCTATATACTTTGATGCTGAGCGGATGCTTTGAACATTTGAACAGCA 185
QY 362 TAATGTTAAAGCTGTGGATGAAGCAACCATCATTCATGACATTTCTAACTAAGGCAAAATG 421
Db 186 TCAAGACCAAGGTTGGATGAGTCAACATTTTGACCAACCGCAGCAATG 245
QY 422 CACAGGCTCAACAGATCAAGCAGCATATCTCCAGGAAACAGGAAAGCCCTGGATGAA 481
Db 246 CACAGAGACAGGATATTCCTTCGCTTACCAGAGAGGACCAAAAGGAACTTGCATCAG 305
QY 482 CACTGAAGAAAGCCCTTACAGGTCACTTTCAGAGGTTGTTTACCTCTGCTAAAAATC 541
Db 306 CACTGAAGTCAAGCTTATCTGCCACCTGGAGCGGTGATTTTGGGCTTATTGAAGACAC 365
QY 542 CAGCGCAATTTGATGCTGATGAACCTTCTGCTGCTGCTGATGAAGGCTTGGAACTGTAAG 601
Db 366 CTGCTCAGTATGAGCTTCTGAGCTAAAGCTTCCATGAAGGGGCTGGGAACCGCAGG 425
QY 602 ATACTCTAATTCAGATTTTGGCATCAAGAACTAAACAAAGAAATCAGAGACATTAACAGG 661
Db 426 ACTCTCTCATTTGAGATCATCTCTCCAGAACCAACAGGAGCTGAGGAAATTAACAGAG 485
QY 662 TCTACAGAGAGAACTGAAGAGATCTGGCCAAAGACATAACCTCAGACACATCTGGAG 721
Db 486 TCTACAGAGAAATGTACAAAGCTGATCTGGAGAGGACATTTATTCGACACATCTGGTG 545
QY 722 ATTTTGGGACGCTTTGCTTTCTCTGTTGAGGGTACCGATCTCAGGA---CTTTGGTG 778
Db 546 ACTTCGCAAGCTGATGTTTGGCTGGCAAGGGTGAAGAGCAGAGATGCTCTGTCA 605
QY 779 TGAATCAAGACTTTGGCTGATTCAGATGCGCAGGCTTGTATGAAGCAGAGAAAGAGAA 838
Db 606 TTGATTATGAACCTGATTGACCAAGATGCTCGGATCTCTATGACGCTGGAGTGAAGAGA 665
QY 839 AGGGGACAGAGTAAACGTTTCAATACCTCTTACCACAGAGAGCTATCCACACTTC 898
Db 666 AAGGAACTGATGTTCCCAAGTGGATCAGCATCATGACCGGAGCGGTGCCCCACTCC 725
QY 899 GCAGAGTGTTCAGAAATACACCAAGTACAGTAAGCATGACATGAACAAAGTTCTGGACC 958
Db 726 AGAAAGTATTTGATAGGTACAAGGTTACAGCCCTATGACATGTTGGAAGCATCAGGA 785
QY 959 TGGAGTTGAAAGGTGACATTTGAGAAATGCTCAGAGCTATCGTGAAGTGGCCACAGCA 1018
Db 786 AAGAGGTTAAAGGAGACCTGGAATGCTTTCTGAACTGTTTCTGAGTGCATTCAGAA 845
QY 1019 AACGAGCTTTCTTTGAGAGAGAGCTTCAAGCCATGAAGGCTGTTGAACTCGCCATA 1078
Db 846 AGCCCTGTATTTTGTGCTGATCGCTGATGCTATGCTCATGAGGCAAGGGGACCGAGATA 905
QY 1079 AGCATTTGATCAGATTTATGTTTCCGTTTCTGAAATTCAGATGAATGATATCAAGCAT 1138
Db 906 AGGTCTGTATCAGATTCATGCTTCCCGCAGTGAAGTGGACATGTTGAAATTTAGTCTG 965
QY 1139 TCTATCAGAGAGTGTATGTTATCTCCCTTTGCCAAGCCATCTGGATGAACCAAGGAG 1198
Db 966 AATTCAAGAGAAAGTACGCAAGTCCCTGTACTATTATTCAGAGAGACATTAAGGGCG 1025
QY 1199 AGTATGAGAAATCTCTGGTGGCTTTTGGAGGAAACCTAAACATTCCTTGAATGCTCT 1258
Db 1026 ACTACCAAGAGAGGCTGCTGTACTCTGTGTGGAGATGACTGAAGCCGCGACACGCGCTG 1085
QY 1259 A 1259
Db 1086 A 1086

RESULT 5

JS-09-023-655-1149
 ; Sequence 1149, Application US/09023655
 ; Patent No. 6607879
 ; GENERAL INFORMATION:
 ; APPLICANT: Cocks, Benjamin G.
 ; APPLICANT: Susan G. Stuart
 ; APPLICANT: Jeffrey J. Sellhamer
 ; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
 ; TITLE OF INVENTION: EXPRESSION
 ; NUMBER OF SEQUENCES: 1508
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
 ; STREET: 3174 PORTER DRIVE
 ; CITY: PALO ALTO
 ; STATE: CALIFORNIA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/023,655
 ; FILING DATE: HEREWITH
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Zeller, Karen J.
 ; REGISTRATION NUMBER: 37,071
 ; REFERENCE/DOCKET NUMBER: PA-0001 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (650) 855-0555
 ; TELEFAX: (650) 845-4166
 ; INFORMATION FOR SEQ ID NO: 1149:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1339 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: GENBANK
 ; CLONE: g186967
 ; JS-09-023-655-1149

Query Match 17.7%; Score 268.4; DB 4; Length 1339;
 Best Local Similarity 55.2%; Pred. No. 5.2e-72;
 Matches 524; Conservative 0; Mismatches 426; Indels 0; Gaps 0;

2y	286	GGTGGTCCCGGATCAGCGGTGAGCCCTATCCCTACCTTCAATCCATCCCTCGGATGCGCT	345
Db	62	GTGGACACCGAGAACAGTAAGAGATATCCAGACTTTAGCCCATCAGTGGATGCTGAA	121
2y	346	GCCTTGATAGGCCCAATAAGTTAAAGGTGTGATGAAGCAACCATCATTCATCTA	405
Db	122	GCTATTTCAGAAAGCAATCAGAGGAATTGGAACCTGATGAGAAATGCTCATCAGCATTCG	181
2y	406	ACTAAGCGMAACATGACAGCGTCAACAGATCAACAGCATATCTCCAGGAAACAGGA	465
Db	182	ACTGAGAGGTCAATGACAGCGGACGCTGATTTTGAAGAAATATCAGCAGCATATGGA	241
2y	466	AAGCCCTTGATGAAACACCTGAAGAAAGCCCTTACAGGTCACTTGAGGAGGTGTTT	525
Db	242	AAGGAGCTGAAAGATGACTTGAAGGGTGATCTCTCTGCGCCACTTTGAGCATCTCATGGTG	301
2y	526	GCTCTGCTAAAACATCCAGCGCAATTTGCTGATCAACTTCGTGCTGCATGAAGGGC	585
Db	302	GCCTTAGTACTCCACCAGCAGTCTTTTGATGCAAGCAGCTTAAAGAAATCCATGAAGGGC	361

RESULT 6

US-08-526-136-1
 ; Sequence 1, Application US/08526136
 ; Patent No. 6107089
 ; GENERAL INFORMATION:
 ; APPLICANT: Towle, Christine A. et al.
 ; TITLE OF INVENTION: ANNEXIN XI
 ; NUMBER OF SEQUENCES: 36
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fish & Richardson
 ; STREET: 225 Franklin Street
 ; CITY: Boston
 ; STATE: Massachusetts
 ; COUNTRY: U.S.A.
 ; ZIP: 02110-2804
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 ; COMPUTER: IBM PS/2 Model 502 or 55SX
 ; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
 ; SOFTWARE: WordPerfect (Version 5.0)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/526,136
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/214,036
 ; FILING DATE:

QY	586	CTTGGAACTGATGAAGATACCTCTAATTGAGATTTTGGCATCAAGAACTAAACAAGAATC	645
Db	362	CGGGGACAAACGAAGATGCCCTTGAATTAATCTTAATACCAAGCAAGCAGCAATG	421
QY	646	AGAGACATTAAACAGGGTCTACAGAGAGAACTGAAGAGAGATCTGGCCAAAGACATAACC	705
Db	422	AAGGATATCTCTCAAGCTATTATACAGATATACAAAGAGAGTCTTGGAGATGACATTAGT	481
QY	706	TACAGACATCTGAGATATTTTCGAAACGCTTTGCTTCTTCTTGAAGGTGACCGATCT	765
Db	482	TCCGAAACATCTGGTGACTTCCGGAAGCTCTGTTGACTTTTGGCAGATGGCAGAGAT	541
QY	766	GAGGACTTTGGTGTGAATGAAGACTTGCCTGATTCAGATGCCAGGGCTTGTATGAAGCA	825
Db	542	GAAAGTCTGAAGTGGATGAGCATCTGGCCAAACAGATGCCAGATCTCTATAAAGCT	601
QY	826	GGAGAAAGGAAAGGGGACAGACGTAACGCTGTCAATACCATCTTACCCAGAGC	885
Db	602	GGTGAGAACAGATGGGGCAGCGATGAAGACAAATTCACCTGAGATCCTGTGTTAAGGAGC	661
QY	886	TATCCACAACTTCGACAGAGTGTTCAGAAATACACCAAGTACAGTAAAGCATGACATGAAC	945
Db	662	TTTCCTCAATTAACAACTAACATTTGATGATACAGAAATATCACCCAAAGGACATTTGTG	721
QY	946	AAAGTTCTGGACCTGGAGTTGAAAGGTGACATTGAGAAATGCCCTCACAGCTATCGTGAAG	1005
Db	722	GACAGCATAAAGGAGAAATTTCTGGGCAATTTTGAAGACTTACTGTGGCCATAGTTAAT	781
QY	1006	TGCGCCACAAACCAAGCTTTCTTTTCAGAGAGCTTTCATCAAGCATCAAGAGGTGTT	1065
Db	782	TGTGTGAGGAACACGCGCGCTTTTACCCGAAAGACTGCAATCGAGCCTTGAAGGGTATT	841
QY	1066	GGAACTCGCCATAAGGCAATTCATCAGGATTTATGTTTCCCGCTTCTGAAATTCACATGAAT	1125
Db	842	GGAACTGATGAGTTTACTCTCAACCGAATAATGTTCCAGATCAGAAATTCACCTTTTG	901
QY	1126	GATATCAAGCATCTATCAGAAATGATGATATCTCCCTTTGCCAAGCCATCCTGGAT	1185
Db	902	GACATTCGAAACAGAGTTCAAGAGCAATTTAGGCTATTCCTCTATTCAGCAATTAATCG	961
QY	1186	GAAACCAAGGAGAGTATGAGAAATCTGGTGGCTCTTTTGTGGAGGAAA	1235
Db	962	GATACTTCTGGAGACTATGAATCACACTCTTAAATCTGTGGTGAGA	1011

Db	1408	AAAGGCTCAACAGGCGCATGAGGGGAGCCGGAAACCAAGACCGGACCTCGATCCGCATCA	1466
Qy	1097	TGGTTTCCGTTCTGAAATTGACATGAATGATATCAAAAGCATTTCTATCAGAAGATGTATG	1156
Db	1468	TGGTGTCTCGCAGCGAGATCGACCTCTCTGGACATCAGACGACAGATATAAGCGGCTGTATG	1527
Qy	1157	GTATCTCCCTTTCCCAAGCCATCTCTGGATGAAACCAAGGAGAGATATGAGAAAAATCCCTGG	1216
Db	1528	GCAAGTGCCTGTACCAAGCATCACGGGAGACACTTCAGGGGATTCACGAGATTCTGC	1587
Qy	1217	TGGCTCTTTGGAGGAAA	1235
Db	1588	TGAAGATCTGTGGTGCAA	1606
RESULT 7			
US-08-526-136-3			
; Sequence 3, Application US/08526136			
; Patent No. 6107089			
; GENERAL INFORMATION:			
; APPLICANT: Towle, Christine A. et al.			
; TITLE OF INVENTION: ANNEXIN XI			
; NUMBER OF SEQUENCES: 36			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Fish & Richardson			
; STREET: 225 Franklin Street			
; CITY: Boston			
; STATE: Massachusetts			
; COUNTRY: U.S.A.			
; ZIP: 02110-2804			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb			
; COMPUTER: IBM PS/2 Model 50Z or 55SX			
; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)			
; SOFTWARE: Wordperfect (Version 5.0)			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/08/526,136			
; FILING DATE:			
; CLASSIFICATION: 435			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: US/08/214,036			
; FILING DATE:			
; APPLICATION NUMBER: 07/837,775			
; FILING DATE: February 13, 1992			
; APPLICATION NUMBER: 07/764,465			
; FILING DATE: September 23, 1991			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Clark, Paul T.			
; REGISTRATION NUMBER: 30,162			
; REFERENCE/DOCKET NUMBER: 00786/099001			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: (617) 542-5070			
; TELEFAX: (617) 542-8906			
; TELEX: 200154			
; INFORMATION FOR SEQ ID NO: 3:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 2311			
; TYPE: nucleic acid			
; STRANDEDNESS: single			
; TOPOLOGY: linear			
US-08-526-136-3			
Query Match 16.5%; Score 250.2; DB 3; Length 2311;			
Best Local Similarity 54.5%; Pred. No. 2.9e-66;			
Matches 501; Conservative 0; Mismatches 418; Indels 0; Gaps 0			
Qy	317	CTACCTTCAATCCATCCCTCGGATGCTGCTCGCTTGCATAGGCCATATATGGTTAAAGGTG	376
Db	694	CTGGCTTTGACCCCTCGGAGATGCTGAAGTCTCTGGGAAGGCCATGAAAGGCGCTTTGGGA	753
Qy	377	TGGATGACGCAACCATCATTTGACATTTCTACTAAGCGAAACAAATGCACAGCGTCAACAGA	436

754	CTGACGAGCAGGCCATCATTTGATCTGCCTCTGGTGTAGTCGTCTCCAAACGAGCAACGACGACA	813
437	TCAAAGCAGCATATCTCCAGGAAACAGGAAAGCCCTCGATGAAACACTGAAGAAAGCC	496
814	TCCTCTGTCGTTCAAGACAGCATATGGGAAGGATTTGATCAAAGATCTGAATCTGAAC	873
497	TTACAGGTGACCTTGAGGAGGTTGTTTTAGCTCTGCTAAAAACTCCAGCGCAATTTGATG	556
874	TGTCAGAAACTTTTGAGAGCAATCTTGCCCTGATGAAGACCCCTGTCTCTTTGACG	933
557	CTGATGAACCTTCGTGTGTCATGAAGGCGCTTGGAACTGATGAAGTACTCTTAATTTGAGA	616
934	CTTATGAGATAAAGGAAGCTATCAAGGGGCGGGCACTGATGAAGCTGCTCTGATCGAGA	993
617	TTTTGGCATCAAGAACTAAACAAAGAAATCAGAGACATTAACAGGCTCTACAGAGAGGAAC	676
994	TCCTGGCTCCCGACGACGAGCAATCGGGAGCTGAACAGAGTCTACAGACAGAAAT	1053
677	TGAAGAGAGATCTGGCCAAAGACATACCTCAGACACATCTGAGATTTTCGAAACGCTT	736
1054	TCAAAAGACCTTGGAGAGGCCATTTGGAGCAGCACTTCAGGGCACTTCCAGCGGCTCC	1113
737	TGCTTTCTCTGTCTAAGGGTGACCGATCTCAGGACTTTGGTGTGAATGAAGACTTGGCTG	796
1114	TCATCTCTCTCTCAGGGAAAACCGGATGAAGCAAAACGTGGACATGACCTTTGTCC	1173
797	ATTCCAGATGCCAGGCCCTTGATGAAGCAGAGAAAGGAGAAAGGGACAGACGTAAACG	856
1174	AGAGAGATGTGCAGGAGCTCTATGCAGCTGGGAGAACCCGCTGGGAACAGATGAGTCCA	1233
857	TGTTCAATACCATCTTACCACACAGAGCTATCCACAACCTTCGCAGAGTGTTCAGAAAT	916
1234	AGTTCAATCGATTCTGTCTCCGGAGCCGGGCCACCTGGTGGCAGTTTTTAAACGAGT	1293
917	ACACCAAGTACAGTAAGCATGACATGAACAAAGTTCTGGACCTGGAGTTCAAAGGTTGACA	976
1294	ATCAGAGGATGACAGGACGTGACATTCAGAAAGAGCATCTCCGGGAGATGTCCGGGACC	1353
977	TTGAGAAATGCCTCACAGCTATCGTGAAGTGCGCCACAAGCAAAACCAAGCTTTCTTTGCAG	1036
1354	TGGAGCAGGGCATGCTGGCTGTGGTGAATGTCTTAAAGATACCCAGCCCTCTTTTGTCTG	1413
1037	AGAGCTTCATCAAGCCATGAAGGCTGTGGAACCTGCCATAGGCAATTCATCAGGATTA	1096
1414	AAAGGCTCAACAGGCCATGAGGGAGCCGGAACCAAAGCCGACCCCTGATCCGCAATCA	1473
1097	TGGTTTCCCGTCTTGAAATTTGACA TGAATGATATCAAAGCATTTCTATCAGAAGATGTATG	1156
1474	TGGTGTCTCGCAGCGAGATCGACCTCTCTGGACATCAGAGCAGAGTATAAGCGCGCTGTATG	1533
1157	GTATCTCCCTTTGCCAGGCCATCTGTGATGAAACCAAAGGAGAGTATGAGAAATCTCTGG	1216
1534	GCAAGTGCCTGTACCAAGCATACCGGAGACACTTCAGGGGATTAACGGAGATTTCTGC	1593
1217	TGGCTCTTTGAGGAAA	1235
1594	TGAAGATCTGTGTGGCAA	1612

RESULT 8

RESULT 8
 US-09-023-655-1190
 ; Sequence 1190, Application US/09023655
 ; Patent No. 6607879
 ; GENERAL INFORMATION:
 ; APPLICANT: Cocks, Benjamin G.
 ; APPLICANT: Susan G. Stuart
 ; APPLICANT: Jeffrey J. Sellhammer
 ; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
 ; TITLE OF INVENTION: EXPRESSION
 ; NUMBER OF SEQUENCES: 1508
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
 ; STREET: 3174 PORTER DRIVE

```

599 GATGCCAGGACCTGTATGAGCTGGAGAGAAATGGGGACAGATGAGGTGAAATTT 658
862 AATACATCCTTACACAGAGAGTATCCAACTTCGAGAGTGTTCAGAAATACACC 921
659 CTAATCTGTCTCTGTTCCGGAACGAAATACCTGTTGCAATGTTTGTGATGAACAAA 718
922 AAGTACAGTAAGCATGACATGAACAAAGTTCTGGACCTGGAGTTGAAAGGTGACATTGAG 981
719 AGGATATCACAGAGATATTCGAAGAGATTAATCTGAACATCTGGTAGCTTGAA 778
982 AATGCTCACAAGTATCGTGAAGTGGCGCCACAAAGCAACAGCTTTCTTTCGAGAGAG 1041
779 GATGCTCTCTGCTGATAGTAAAGTGCATGAGGAACAAATCTGCATATTTTCTGCAAAAG 838
1042 CTTCTACAGCCATGAAGGTGTGAACTGCCATTAAGGCAATCATCAGGATATGTT 1101
839 CTTCTAATTCAGTAGAGGCTTGGCCACCGATATACACCTCATCAGAGTATGTT 898
1102 TCCGTTCTGAATGACATGAATGATATCAAGCATTTCTATCAGAAGATGATGTTATC 1161
899 TCTCGAGCAGAAATGACATGTTGGATATCCGGGCACACTTCAAGAGACTCTATGGAAG 958
1162 TCCCTTTGCCAGCCATCTCGATGAACCAAGAGAGATGAGAAATCTCTGTGCT 1221
959 TCTCTGATCTGTTTCAAGGGTGACATCTGGAGACTACAGGAAGTACTGCTTGT 1018
1222 CTTTGTGGAGAAA 1235
1019 CTTGTGGAGGAGA 1032

```

RESULT 9

```

US-08-125-746-4
; Sequence 4, Application US/08125746
; Patent No. 5591633
; GENERAL INFORMATION:
; APPLICANT: SAINO, YUSHI
; APPLICANT: IWASAKI, AKIO
; APPLICANT: SUDA, MAKOTO
; TITLE OF INVENTION: ANTICOAGULANT POLYPEPTIDE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/125,746
; FILING DATE: 24-SEP-1993
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/807,623
; FILING DATE: 13-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 037227/1987
; FILING DATE: 20-FEB-1987
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 184428/1987
; FILING DATE: 23-JUL-1987
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5591633man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 80-074-0 DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000

```

```

; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 957 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; US-08-125-746-4

Query Match 15.8%; Score 239.8; DB 1; Length 957;
Best Local Similarity 53.3%; Pred. No. 2.5e-63;
Matches 505; Conservative 0; Mismatches 442; Indels 0; Gaps 0;

289 GGTCCCGGATCAGCGGTGAGCCCTATCTACCTTCAATCCATCTCTCGGATGTCGCTGCC 348
7 GTTCTCAGAGGCACTGTGACTGCTTCCCTGGATTGATGAGCGGCTGATGACAGAACT 66
349 TTGCATAAGGCCATATGTTTAAAGTGTGGATGAAGCAACCATCATTCATTCACATTTCACT 408
67 CTTGCGAAGGCTATGAAAGGCTTTGGGCACAGATGAGGAGCATCTCTGATCTCTGTTGACA 126
409 AAGCGAAACAATCCACAGCGTCAACAGATCAAAGCAGCATATCTCCAGGAAACAGGAAAG 468
127 TCCGAAAGTATGCTCAGCGCCAGGAAATCTCTGCACTTTTAAGACTCTGTTTGCAGG 186
469 CCGCTGGATGAACACTGAAGAAAGCCCTTACAGTCACTCTGAGGAGGTTGTTTAGCT 528
187 GATCTTCTGGATGACCTGAAATCAGAACTAACTGGAAAAATTTGAAAAATTTAATTGGGCT 246
529 CTGCTAAAAACTCCAGCGCAATTTGATGCTGATGAATCTCGTGCTGCCATGAAGGCGCTT 588
247 CTGATGAACCCCTCTCGGCTTTATGATGCTTATGAACTGAACATGCTTGAAGGAGCT 306
589 GGAACCTGATGAAGATACTCTAATTGAGATTTTGGCATCAAGAACTAAACAAAGAAATCAGA 648
307 GGAACAAATGAAAAAGTACTGACAGAAATTTATGCTTCAAGGACACCTGAGAACTGAGA 366
649 GACATTAACAGGCTCTACAGAGAGGAACTGAAGAGAGATCTGGCCAAAGACATTAACCTCA 708
367 GCATCAACAAGTTTATGAAAGAAATATGGCTCAAGCTGGAAGATGACGTGGTGGGG 426
709 GACACATCTGGAGATTTTCGAAACGCTTTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 768
427 GACACTTCAGGCTACTACAGCGGATGTTGGTGGTCTCTCTCTCAGGCTAACAGAGACCCT 486
769 GACTTTGGTGTGAATGAGACTTTGGCTGATTCAGATGCCAGGCGCTTGTATGAAGCAGGA 828
487 GATGCTGGAATGATGAAGCTCAAGTTGAACAAGATGCTCAGGCTTTTATTCAGGCTGGA 546
829 GAAAGGAGAAAGGGGACAGACGTAACGCTGTTTCAATACCATCTCTTACCACCAAGCTAT 888
547 GACTTAAATGGGGACAGATGAAGAAAGTTTATCACCATCTTTGGAACACCAAGCTGTG 606
889 CCACAACTTCGAGAGTGTTCAGAAATACACCAAGTACAGTAAAGCATGACATGAACAAA 948
607 TCTCAITTTGAGAAAGGTGTTTGAACAAGTACATGACTATATCAGGATTTCAATTTAGGAA 666
949 GTTCTGACCTGGAGTTGAAAGGTGACATTTGAGAAATGCTTCACAGCTATCGTGAAGTGC 1008
667 ACCATTGACCGGAGACTTCTGGCAATTTAGAGCAACTACTCTCTGCTGTTGTGAAATCT 726
1009 GCCACAAAGCAACAGCTTTCTTTTGGAGAGAGGTTTCAATCAAGCCATGAAGGTGTGGA 1068
727 ATTCGAAAGTATACCTGCTACCTTGCAGAGACCTCTTATTTATCTATGAAGGAGGCTGGG 786
1069 ACTCGCATTAAGGCAATGATCAGGATTTATGTTTCCGTTTCTGAAATGACATGAATGAT 1128
787 ACAGATGATCATACCTCTCATCAGAGTATGTTTCCAGGAGTGAATGATGATCTGTTTAA 846
1129 ATCAAGCATTTCTATCAGAGAGATGATGTTATCTCCCTTTGCCAAGCCATCTCGATGAA 1188

```

Db 847 ATCAGGAAGAGGATTAGGAAGAAATTTTCCACCTCTCTTTATTCATGATTAAGGGAGAT 906
Qy 1189 ACCAAAGAGAGATGAGAAATCTCGTGCTCTTTGTGGAGAAA 1235
Db 907 ACATCTGGGACTATAGAAAGCTCTTCTGCTGCTGTGGAGA 953
RESULT 10
522537-3
;PATENT No. 522537
;APPLICANT: FOSTER, DONALD
;TITLE OF INVENTION: METHODS FOR PRODUCING HYBRID
;PHOSPHOLIPID-BINDING PROTEINS
;NUMBER OF SEQUENCES: 14
;CURRENT APPLICATION DATA:
;APPLICATION NUMBER: US/07/459,082
;FILING DATE: 29-DEC-1989
;SEQ ID NO:3
;LENGTH: 1460
522537-3
Query Match 15.8%; Score 239.8; DB 6; Length 1460;
Best Local Similarity 53.3%; Pred. No. 3.3e-63;
Matches 505; Conservative 0; Mismatches 442; Indels 0; Gaps 0;
Qy 289 GGTCCCGATCAGCGGTGAGCCCTATCTACCTTCAATCCATCCTCGGATGCGCTGCC 348
Db 22 GTTCTCAGAGGACGTGACTGACTTCTCCTGGATTGATGAGCGGCTGATGAGAACT 81
Qy 349 TTGCATAAGGCATTAATGGTTAAAGGTGGATGAAGCAACCATCATGACATTTCACT 408
Db 82 CTTCCGAAGGCTATGAAAGGCTTGGCCACAGATGAGGAGCATCTGACTCTCTTGACA 141
Qy 409 AAGCGAACAATGACAGCGTCAACAGATCAACAGCATATCTCCAGGAACAGGAAG 468
Db 142 TCCGAAGTAAATGCTCAGCGCCAGGAATCTCTGAGCTTTTAAGACTCTGTTGGCAGG 201
Qy 469 CCCCTGGATGAACACTGAAGAAACCCCTTACAGTCACTTGGAGAGGTTGTTTACGT 528
Db 202 GATCTTCTGGATGACCTGAATCAGAACTAACTGGAAATTTGAAATTAATTTGGCT 261
Qy 529 CTGCTAAACCTCAGCGCAATTTGATCTGATGACTTCTGCTGCTGCAAGGCGCTT 588
Db 262 CTGATGAACCCCTCGCGCTTTATGCTTATGACTTGAACATGCAAGCTGGAAGGAGCT 321
Qy 589 GGAATGATGAAGTACTTAATGAGATTTGGCATCAAGAACTTAAACAAAGAAATCAGA 648
Db 322 GGAACAAATGAANAAGTACTGACAGAAATTTGCTTCAAGGACACCTGAAGAACTGGA 381
Qy 649 GACATTAACAGGCTTACAGAGAGAACTGAAGAGATCTGGCCAAAGACATTAACCTCA 708
Db 382 GCCATCAACAAAGTTTATGAAGAAGAAATATGCTCAAGCCTGGAAGATGACGTGGTGGG 441
Qy 709 GACATCTGGAGATTTTCGGAACGCTTTGCTTCTTCTGTAAGGTTGACCGATCTGAG 768
Db 442 GACATTCAGGTTACTACAGCGGATGTTGGTGTCTCTCTCAGCTTAACAGAGCCCT 501
Qy 769 GACTTTGGTGAATGAAGACTTTGCTGATTCAGATGCGGCGCTGATGAAGACGAGA 828
Db 502 GATGCTGGAATTTGAAAGCTCAAGTTGAACAAGATGCTCAGGCTTTATTTTCAGGCTGA 561
Qy 829 GAAAGGGAAGGGGACAGCACTTAACTGTTCAATACCTCCCTTACCACAGAGCTAT 888
Db 562 GAATTTAATGGGGACAGATGAAGAAAGTTTATCAACATCTTTGGAAACAGAAAGTGTG 621
Qy 889 CCACAACCTTCGAGAGTGTTCAGAAATACACCAAGTACAGTAAGCATGACATGAACAAA 948
Db 622 TCTCATTTTGAAGAAAGTGTTCACAAGTACATGACTATATCAGGATTTCAAATTTAGGAA 681
Qy 949 GTTCTGGACTGGATTGAAGGTCATGAGAAATGCTTCAAGCTATCGTGAAGTGC 1008
Db 682 ACCATTGACCGGAGACTTCTGGCAATTTAGAGCAACTACTCTTGTGTTGAAATCT 741

Qy 1009 GCCACAAGCAACACAGCTTTCTTTCAGAGAGCTTCAATCAAGCCATGAAAGCTTTGGA 1068
Db 742 ATTCAAGTATATCTGCTGCTACCTTTCAGAGACCCCTCTATTATGCTATGAAGGAGCTGG 801
Qy 1069 ACTCGCCATTAAGGCATTTGATCAGGATTTATGGTTTCCCGTCTTGAATTCACATGAATGAT 1128
Db 802 ACAGATGATCATACCTCATCAGACTCATGTTTCCAGAGTGAGATTGATCTGTTTAAAC 861
Qy 1129 ATCAAAGCAATTCATCAGAGATGATGATCTCCCTTTCGCAAGCCATCTCTGATGAA 1188
Db 862 ATCAGGAAGGAGTTTAGGAAGAAATTTGCCACCTCTCTTTATTCATGATTAAGGAGAT 921
Qy 1189 ACCAAAGGAGATGATGAGAAATCTGTTGGTCTTTCTGTGGAGAAA 1235
Db 922 ACATCTGGGACTATAGAAAGCTTTCTGCTGCTCTGTGGAGAAGA 968
RESULT 11
US-08-125-746-2
; Sequence 2, Application US/08125746
; Patent No. 5591633
; GENERAL INFORMATION:
; APPLICANT: SAINO, YUSHI
; APPLICANT: IWASAKI, AKIO
; APPLICANT: SUDA, MAKOTO
; TITLE OF INVENTION: ANTICOAGULANT POLYPEPTIDE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/125,746
; FILING DATE: 24-SEP-1993
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/807,623
; FILING DATE: 13-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 037227/1987
; FILING DATE: 20-FEB-1987
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 184428/1987
; FILING DATE: 23-JUL-1987
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5591633man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 80-074-0 DIV
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1567 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 136..1095
US-08-125-746-2

Query Match 15.8%; Score 239.8; DB 1; Length 1567;
 Best Local Similarity 53.3%; Pred No. 3.5e-63;
 Matches 505; Conservative 0; Mismatches 442; Indels 0; Gaps 0;

289 GTTCCGGATCAGCGGTGAGCCCTTATCTACCTTCAATCCATCTCGATGTCGTGCC 348
 145 GTTCTCAGAGGCACTGTGACTGACTTCCCTGGAATTTGATGAGCGGCTGTGCGAAACT 204
 349 TTGCATAGCCCATATGTTAAAGGTGGATGAAGCAACCATCATTTGACATTTCACT 408
 205 CTTGGAAGGCTATGAAGGCTTGGCAGATGAGGAGCATCTGACTCTGTGACA 264
 409 AAGCGAAACAATGCAAGCGTCAACAGATCAAGAGCATATCTCCAGAAACAGAAAG 468
 265 TCCGAAGTAATGCTCAGCGCCAGGAATCTCTGCACTTTTAAGACTCTGTTGGCAGG 324
 469 CCCCTGGATGAACACTGAAGAACCCCTTACAGTCAACCTGAGGAGGTTGTTTACCT 528
 325 GATCTCTCGATGACCTGAAATCAGAACTAACTGGAATTTGAAATTAATTTGCTGCT 384
 529 CTGCTAAAACCTCCAGCGCAATTTGATGCTGATGAACCTTGTGCTGCCATGAAGGCTT 588
 385 CTGATGAACCCCTCTCGGCTTTATGATGCTTATGAATGAACATGAAACATGCTTGAAGGAGCT 444
 589 GGAACGTGATGATGACTCTAATTTGATGATTTGGCATCAAGACTTAAACAAGATCAGA 648
 445 GGAACAAATGAAGAAAGTACTGACAGAAATTAATGCTTCAAGGACACCTGAAGACTGAGA 504
 649 GACATTAACAGGGTCTACAGAGAGGAATGAAGAGAGATCTGCGCCAAAGACATACTCA 708
 505 GCATCAACAAGTTTATGAAGAAGATATGGCTCAAGCTGGAAGATGACGTGGTGGG 564
 709 GACATCTGGAGATTTTGGAAAGCTTTGCTTTCTTTGCTAAGGGTACCGATCTGAG 768
 565 GACACTTACGGGTACTACAGCGGATGTTGGTGGTTCTCTTCAAGCTTAAAGACCCCT 624
 769 GACTTTGGTGTGAATGAAGACTTTGGCTGATTCAGATGCCAGGCGCTTGTATGAAGCAGA 828
 625 GATGCTGGAATTTGATGAAGCTCAAGTTGAACAGATGCTCAGGCTTTATTTTCAAGCTGGA 684
 829 GAAAGAGAAAGGGGACAGAGTAAAGCTGTTTCAATACCATCTTACACAGAGAGCTAT 888
 685 GAACTTTAAATGGGGGACAGATGAAGAAAGTTTATCACCCTCTTTGGAACAGAAAGTGTG 744
 889 CCACAACTTCGAGAGTGTGTTTCAAGAAATACACCAAGTACAGTAAAGCATGACATGAACAAA 948
 745 TCTCATTTGGAAGGTTGTTGACAGTACATGACTATATCAGGATTTCAATTTGAGGAA 804
 949 GTTCTGGAACCTGAGTTGAAGGTGACATGAGAAATGCTCAGACTATCTGTAAGTGC 1008
 805 ACCATTGACCGGAGACTTCTGGCAATTTAGAGCAACTACTCTCTGCTGTTTGAATCT 864
 1009 GCCACAGCAACCCAGCTTTCTTTGAGAGAGCTTTCATCAAGCCATGAAGGTGTTGA 1068
 865 ATTCGAAGTATACCTGCTACCTTTGAGAGACCCCTCTATTTGCTATGAAGGAGCTGG 924
 1069 ACTCGCATAGGCAATGATCAGATTAATGTTTCCGTTCTGAAATTTGACATGAATGAT 1128
 925 ACAGATGATCATACCTCTCATCAGAGTCACTGTTTCCAGGAGTGAATGATCTGTTTAA 984
 1129 ATCAAGCATCTATCAGAGATGATGTTATCTCCCTTTGCCAGCCATCTCGATGAA 1188
 985 ATCAGGAAGGAGTTTGAAGAAGATTTTGGCACTCTCTTTTATTCATGATTAAGGAGAT 1044
 1189 ACCAAGGAGATGATGAGAAATCTCTGGTGGCTCTTTTGTGGAGAAA 1235
 1045 ACATCTGGGACTATAAGAAAGCTCTTCTGCTGCTCTGTTGAGAGA 1091

RESULT 12
 US-09-324-096A-1
 ; Sequence 1, Application US/09324096A
 ; Patent No. 6323313

GENERAL INFORMATION:
 ; APPLICANT: Tait, Jonathan
 ; TITLE OF INVENTION: ANNEXIN DERIVATIVE WITH ENDOGENOUS CHELATION SITES
 ; FILE REFERENCE: US/09-324-096A
 ; CURRENT APPLICATION NUMBER: US/09-324-096A
 ; NUMBER OF SEQ ID NOS: 12
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 1
 ; LENGTH: 981
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(981)
 ; US-09-324-096A-1

Query Match 15.6%; Score 236.6; DB 4; Length 981;
 Best Local Similarity 53.1%; Pred No. 2.5e-62;
 Matches 503; Conservative 0; Mismatches 444; Indels 0; Gaps 0;

289 GGTCCGGATCAGCGGTGAGCCCTTATCTTCAATCCATCTCGATGTCGTGCC 348
 31 GTTCTCAGAGGCACTGTGACTGACTTCCCTGGAATTTGATGAGCGGCTGTGCGAAACT 90
 349 TTGCATAGGCCCATATGTTAAAGGTGGATGAAGCAACCATCATTTGACATTTCAACT 408
 91 CTTGGAAGGCTTATGAAGGCTTGGGACAGATGAGGAGCATCTGACTCTGTGACA 150
 409 AAGCGAAACAATGCAAGCGTCAACAGATCAAGAGAGCATATCTCCAGGAAACAGAAAG 468
 151 TCCGGAAGTAAATGCTCAGCGCCAGGAATCTCTGCACTTTTAAAGACTCTGTGTCAGG 210
 469 CCCCTGATGAACACTGGAAGAGCCCTTACAGGTCACCTTCAAGAGGTTGTTTAGCT 528
 211 GATCTTCTGATGACCTGGAATCAGAACTAACTTGAAGAAATTTGAAATTAATTTGGCT 270
 529 CTGCTAAAAACTCCAGCGCAATTTGATGCTGATGAACCTTCTGCTCCATGAAGGCGCTT 588
 271 CTGATGAACCCCTCTCGGCTTTATGATGCTTATGAACATGAACATGCTTTGAAGGAGCT 330
 589 GGAATCATGAGATCTCTAATTTGATGATTTTGGCATCAAGAACTAAACAAGAAATCAGA 648
 331 GGAACAAATGAAGAAAGTACTGACAGAAATTTATGCTTCAAGGACACCTGGAAGACTGAGA 390
 649 GACATTAACAGGGTCTACAGAGAGGAACCTGAAGAGAGATCTGCGCAAGACATAAOCCTCA 708
 391 GCCATCAACAAGTTTATGAAGAGATATGGCTCAAGCTTGAAGATGACGTGTTGGG 450
 709 GACATCTGAGATTTTGGAAAGCTTTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 768
 451 GACACTTCAAGGCTACTACAGCGGATGTTGGTGGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 510
 769 GACTTTGGTGTGAATGAAGACTTGGCTGATTCAGATGCCAGGCGCTTGTATGAAGCAGGA 828
 511 GATCTGGAATTTGATGAGCTCAAGTTGAACAGATGCTCAGGCTTTATTTTCAAGCTGGA 570
 829 GAAAGGAGAAAGGGAACAGAGTAAAGCTGTTTCAATACCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 888
 571 GAACTTAAATGGGGGACAGATGAAGAAAGTTTATCACCCTTTTGGAAACAGAAAGTGTG 630
 889 CCACAACTTCCAGAGGTTTTCAGAAATACACCAAGTACAGTAAGCATGACATGAACAAA 948
 631 TCTCATTTGAGAAAGGTTTGGACAGTACATGACTATATCAGGATTTCAATTTGAGAA 690
 949 GTTCTGGAACCTGGAAGTGAAGGAGTGAATTTGAGAAATGCTTCAAGATATCTGTTGAGTGC 1008
 691 ACCATTGACCGGAGACTTCTGGCAATTTTAGAGCAACTACTCTTCTTCTTCTTCTTCTTCTTCTTCT 750
 1009 GCCACAAACCAAGCTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1068
 751 ATTGGAATATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 810

3603 CTATATGCAATTAACCTGTTTGTGAACCTCCGATCTTCCTTTCCGATACCTATTGAT 3662
 128867 CTATATGCAATTAACCTGTTTGTGAACCTCCGATCTTCCTTTCCGATACCTATTGAT 128808
 3663 GAATAAGAAATTAACCTGA 3682
 128807 GAATAAGAAATTAACCTGA 128788

RESULT 5
 AC068844 170233 bp DNA linear HTG 24-AUG-2002
 LOCUS Homo sapiens chromosome 3, clone RP11-690F8 map 3, WORKING DRAFT
 DEFINITION SEQUENCE, 23 unordered pieces.
 AC068844
 VERSION AC068844.2 GI:8516083
 KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 170233)
 Birren,B., Linton,L., Nusbaum,C. and Lander,E.
 Homo sapiens chromosome 3, clone RP11-690F8
 Unpublished
 2 (bases 1 to 170233)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
 Anderson,S., Baldwin,J., Barna,N., Bastien,V., Bada,F.,
 Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
 Campopiano,A., Castile,A., Choepel,Y., Colangelo,M., Collins,S.,
 Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
 Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
 Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
 Grand-pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
 Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
 Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczyk,J.,
 Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
 McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
 Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mieng,V., Morrow,J.,
 Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
 O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
 O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
 Pisan,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
 Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
 Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talanas,J.,
 Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
 Vassiliev,H., Viel,R., Vo.A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
 Young,G., Zainoun,J., Zimmer,A. and Zody,M.

Direct Submission
 Submitted (10-MAY-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 170233)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
 Anderson,S., Baldwin,J., Barna,N., Bastien,V., Bada,F.,
 Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
 Campopiano,A., Castile,A., Choepel,Y., Colangelo,M., Collins,S.,
 Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
 Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
 Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
 Grand-pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
 Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
 Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczyk,J.,
 Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
 McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
 Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mieng,V., Morrow,J.,
 Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
 O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
 O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
 Pisan,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
 Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
 Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talanas,J.,
 Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
 Vassiliev,H., Viel,R., Vo.A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
 Young,G., Zainoun,J., Zimmer,A. and Zody,M.

Direct Submission
 Submitted (10-MAY-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 170233)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
 Anderson,S., Baldwin,J., Barna,N., Bastien,V., Bada,F.,
 Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
 Campopiano,A., Castile,A., Choepel,Y., Colangelo,M., Collins,S.,
 Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
 Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
 Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
 Grand-pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
 Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
 Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczyk,J.,
 Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
 McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
 Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mieng,V., Morrow,J.,
 Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
 O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
 O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
 Pisan,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
 Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
 Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talanas,J.,
 Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
 Vassiliev,H., Viel,R., Vo.A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
 Young,G., Zainoun,J., Zimmer,A. and Zody,M.

TITLE

Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Jun 14, 2000 this sequence version replaced GI:7767838.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu

Center project name: L7518
 Center clone name: 690 F 8

Sequencing vector: M13: M77815; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731

Consensus quality: 156993 bases at least Q40
 Consensus quality: 163905 bases at least Q30
 Consensus quality: 166725 bases at least Q20

Insert size: 244000; agarose-fp
 Insert size: 168033; sum-of-contigs
 Quality coverage: 3.8 in Q20 bases; agarose-fp
 Quality coverage: 5.6 in Q20 bases; sum-of-contigs

***** NOTE: This is a 'working draft' sequence. It currently
 consists of 23 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.

1 1957: contig of 1957 bp in length
 1958 2057: gap of 100 bp
 2058 3880: contig of 1823 bp in length
 3881 3980: gap of 100 bp
 3981 5189: contig of 1209 bp in length
 5190 5289: gap of 100 bp
 5290 7410: contig of 2121 bp in length
 7411 7510: gap of 100 bp
 7511 8754: contig of 1244 bp in length
 8755 8954: gap of 100 bp
 8955 10300: contig of 2046 bp in length
 10301 11000: gap of 100 bp
 11001 14682: contig of 3682 bp in length
 14683 14782: gap of 100 bp
 14783 19532: contig of 4750 bp in length
 19533 19632: gap of 100 bp
 19633 25241: contig of 5609 bp in length
 25242 25341: gap of 100 bp
 25342 32065: contig of 6724 bp in length
 32066 32165: gap of 100 bp
 32166 40193: contig of 8028 bp in length
 40194 40293: gap of 100 bp
 40294 46461: contig of 6168 bp in length
 46462 46561: gap of 100 bp
 46562 54584: contig of 8093 bp in length
 54585 54754: gap of 100 bp
 54755 63205: contig of 8451 bp in length
 63206 69683: contig of 6378 bp in length
 69684 69783: gap of 100 bp
 69784 71773: contig of 7390 bp in length
 71774 72773: gap of 100 bp
 72774 86740: contig of 9467 bp in length
 86741 86840: gap of 100 bp
 86841 94331: contig of 7491 bp in length
 94332 94431: gap of 100 bp
 94432 101503: contig of 7072 bp in length
 101504 101603: gap of 100 bp
 101604 118443: contig of 16840 bp in length

2y	2103	GCTCATTGCACTAACTAGAGGGTGAGTTGGCCCTGTCTTCTCTTTTTTCTCGAACCTCAG	2162
Db	150424	GCTCATTTGCACATAACTAGAGGGTGAGTTGGCCCTGTCTTCTTTTTTCTCGAACCTCAG	150483
2y	2163	TTTTCCTCAGTGAAGTAAAGAATGCACTAACCCTTTTGAATTTGATAAGTTATAAAATCTG	2222
Db	150484	TTTTCCTCAGTGAAGTAAAGAATGCACTAACCCTTTTGAATTTGATAAGTTATAAAATCTG	150543
2y	2223	TGTTCTTGATCAATTTGGTCCAGAGGGGAGATAGGTTCCCTGTGATTTTTTCTTCTTCTAT	2282
Db	150544	TGTTCTTGATCAATTTGGTCCAGAGGGGAGATAGGTTCCCTGTGATTTTTTCTTCTTCTAT	150603
2y	2283	AGAAATAAAATGAAATCTTGTACTAGAACAGAAATGTCCAGATGGCCAAAACAAGATGAC	2342
Db	150604	AGAAATAAAATGAAATCTTGTACTAGAACAGAAATGTCCAGATGGCCAAAACAAGATGAC	150663
2y	2343	CAGATTTGATCTCACGCTGATGACCTACAGGTCTGTCTATGATATGGAGTCTCATGG	2402
Db	150664	CAGATTTGATCTCACGCTGATGACCTACAGGTCTGTCTATGATATGGAGTCTCATGG	150723
2y	2403	TAAAGCAGAAGAGAGTGGGAAGAGACCAACCCACCTCTGTCTTTCATATTTGCCATTCCA	2462
Db	150724	TAAAGCAGAAGAGAGTGGGAAGAGACCAACCCACCTCTGTCTTTCATATTTGCCATTCCA	150783
2y	2463	TGTTTTAACCTCCGGCTGGAAATAGAAAAGCATTCCTCTTAGAGATGAGGATAAAAAAGAT	2522
Db	150784	TGTTTTAACCTCCGGCTGGAAATAGAAAAGCATTCCTCTTAGAGATGAGGATAAAAAAGAT	150843
2y	2523	TCAGATTCAAAGGGGGAGAAAATGGAGATTTAATCTCTAAACTGTGACTTGGGGAGT	2582
Db	150844	TCAGATTCAAAGGGGGAGAAAATGGAGATTTAATCTCTAAACTGTGACTTGGGGAGT	150903
2y	2583	CAGTCAATTTACAGTTAGTCTGTGCTTTCCAGCTTCTGTGATTTAAACCCCACTCACTA	2642
Db	150904	CAGTCAATTTACAGTTAGTCTGTGCTTTCCAGCTTCTGTGATTTAAACCCCACTCACTA	150963
2y	2643	CCCTGTTTCAGATGCAATTTGGAAATACCAAAGATTAATTCCTTGBACATPAAGATCTCATTTG	2702
Db	150964	CCCTGTTTCAGATGCAATTTGGAAATACCAAAGATTAATTCCTTGBACATPAAGATCTCATTTG	151023
2y	2703	CAGAAAGCAGATTAAGACCATCAGAAGGAAATTTATTTAGTTGTAATGCACAGGCAACT	2762
Db	151024	CAGAAAGCAGATTAAGACCATCAGAAGGAAATTTATTTAGTTGTAATGCACAGGCAACT	151083
Qy	2763	GTGAGAAACTGTGTGCCAAAATAAGAAATTCCTCTAGTTTTCTTGTCTCTCATTTGAAA	2822
Db	151084	GTGAGAAACTGTGTGCCAAAATAAGAAATTCCTCTAGTTTTCTTGTCTCTCATTTGAAA	151143
Qy	2823	GGAGAAAATTCACATTTGTTTAGCAATTTCAAGCTTTTATGTAATCCATCCCATCTAAAAAC	2882
Db	151144	GGAGAAAATTCACATTTGTTTAGCAATTTCAAGCTTTTATGTAATCCATCCCATCTAAAAAC	151203
2y	2883	TCFTTCAAACTCCACATGTTTCAGTCTGAAATGCAGTCCCTGTCCAAGTGCTTTGGAGAAC	2942
Db	151204	TCFTTCAAACTCCACATGTTTCAGTCTGAAATGCAGTCCCTGTCCAAGTGCTTTGGAGAAC	151263
Qy	2943	TCACAGCAGACGCCTTAATCAAAGTTTTTACAGCCCTTGGACACATATGGGAGGAGGC	3002
Db	151264	TCACAGCAGACGCCTTAATCAAAGTTTTTACAGCCCTTGGACACATATGGGAGGAGGC	151323
Qy	3003	AAGAGTACACCAATTTGTTTAAAGCAAGAAACCAAGTGTCTTTCCTACTAGTCAATTTAGA	3062
Db	151324	AAGAGTACACCAATTTGTTTAAAGCAAGAAACCAAGTGTCTTTCCTACTAGTCAATTTAGA	151383
Qy	3063	ACATGTTATTCATCCAAGACTACTCTACCCCTGCAACATTTGAATCTCCAAGAGCAAAATCCA	3122
Db	151384	ACATGTTATTCATCCAAGACTACTCTACCCCTGCAACATTTGAATCTCCAAGAGCAAAATCCA	151443
Qy	3123	CATTCTCTTAGTCTGAGCTTCTGTGTAATAGGGCAGCTGTCGTCTATGCCCTAGA	3182
Db	151444	CATTCTCTTAGTCTGAGCTTCTGTGTAATAGGGCAGCTGTCGTCTATGCCCTAGA	151503
Qy	3183	ATCACATGATCTGAGGACCAATTCATGGAAGCTGTGTAATAGCCCTAGTCTGGGAGTCTTC	3242

Db	151504	ATCACATGATCTGAGGACCAATTCATGGAAGCTGCTAAATAGCCTAGTCTGGGAGTCTTC	151563
Qy	3243	CATAAAGTTTTCATCTGGAGCAAAACAAACAGGATTAATACTAGGTTTGTTCTCTTCAGCCCT	3302
Db	151564	CATAAAGTTTTCATCTGGAGCAAAACAAACAGGATTAATACTAGGTTTGTTCTCTTCAGCCCT	151623
Qy	3303	CTAAAGAGCATAGGGCTTAGCTGACGAGGCTTCCTTGGGCTTCTCTGTGTGTAGTTTGG	3362
Db	151624	CTAAAGAGCATAGGGCTTAGCTGACGAGGCTTCCTTGGGCTTCTCTGTGTGTAGTTTGG	151683
Qy	3363	TAAACACTATAGCATCTGTTAAGATCCAGTGTCCATGGAAACCTTCCACATGCCGTGAC	3422
Db	151684	TAAACACTATAGCATCTGTTAAGATCCAGTGTCCATGGAAACCTTCCACATGCCGTGAC	151743
Qy	3423	TCGTGGACTATACGTTTTTGGAAAGCAGGGTTCTCTGCTGCTTAACAAGCCACAGTGG	3482
Db	151744	TCGTGGACTATACGTTTTTGGAAAGCAGGGTTCTCTGCTGCTTAACAAGCCACAGTGG	151803
Qy	3483	ACCAGTCTCAATGTCTTCCCTTTACACCTATGTTTTTAAGTAGTCAAACTTCAAGAAACA	3542
Db	151804	ACCAGTCTCAATGTCTTCCCTTTACACCTATGTTTTTAAGTAGTCAAACTTCAAGAAACA	151863
Qy	3543	ATCTAAACAAGTTCTGTGTGCATATGTTTTGTGAACTGTGATTTGTATTGTAGTAGCCTT	3602
Db	151864	ATCTAAACAAGTTCTGTGTGCATATGTTTTGTGAACTGTGATTTGTATTGTAGTAGCCTT	151923
Qy	3603	CTATATTGCATTTAACTTGTTTTTTGTAACCTCCTGATCTCTCTTTCGATACCTATTGAT	3662
Db	151924	CTATATTGCATTTAACTTGTTTTTTGTAACCTCCTGATCTCTCTTTCGATACCTATTGAT	151983
Qy	3663	GAATAAAGAAATTAAGTGA 3682	
Db	151984	GAATAAAGAAATTAAGTGA 152003	
RESULT 6			
AF288537			
LOCUS	AF288537	2180 bp mRNA linear PRI 04-FEB-2001	
DEFINITION	Homo sapiens follistatin-related protein mRNA, partial cds.		
ACCESSION	AF288537		
VERSION	AF288537.1	GI:12658308	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1 (bases 1 to 2180)		
AUTHORS	Vertegaal, A.C., Kuiperij, H.B., van Laar, T., Scharnhorst, V., van der Eb, A.J., and Zantema, A.		
TITLE	cDNA micro array identification of a gene differentially expressed in adenovirus type 5- versus type 12-transformed cells		
JOURNAL	FEBS Lett. 487 (2), 151-155 (2000)		
MEDLINE	20583630		
PUBMED	11150499		
REFERENCE	2 (bases 1 to 2180)		
AUTHORS	Vertegaal, A.C., Kuiperij, B. and Zantema, A.		
TITLE	Direct Submission		
JOURNAL	Submitted (20-JUL-2000) Molecular Cell Biology, Leiden University Medical Center, Wassenaarsweg 72, Leiden 2333 AL, Netherlands		
FEATURES			
source	1. .2180		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	<1. .382		
	/codon_start=2		
	/product="follistatin-related protein"		
	/protein_id="AAK01083.1"		
	/db_xref="GI:12658303"		
	/translation="TTPDQNNKLLRGICVDALIELSDENADWKLQFQELKCLNPSF NPPEKKALEDETADGETEDVDCNVCVCGANNVCTAMTCDGNKQAGTQTTEEMT RVVQELQKEQTAETKREVTKEI"		
CDS			

ORIGIN

Query Match 56.6%; Score 2089; DB 9; Length 2180;
 Best Local Similarity 99.2%; Pred. No. 0;
 Matches 2163; Conservative 0; Mismatches 10; Indels 8; Gaps 6;

QY 628 AACGATCCAGACCGAGAGAACACAAAGTGTAGGGGACTCTGTGTGATGCTTCAT 687
 DB 1 AACGATCCAGACCGAGAGAACACAAAGTGTAGGGGACTCTGTGTGATGCTTCAT 60
 QY 688 TGAACGTCTGATGAAAATGCTGATGAAAATCTCAGCTTCCAGAGTTTCTCAAGTCCCT 747
 DB 61 TGAACGTCTGATGAAAATGCTGATGAAAATCTCAGCTTCCAGAGTTTCTCAAGTCCCT 120
 QY 748 CAACCCATTTTCCAAACCTCTCTGAGAGAGAGTGTGCTTCCAGAGTAAAGCTATGACAG 807
 DB 121 CAACCCATTTTCCAAACCTCTCTGAGAGAGAGTGTGCTTCCAGAGTAAAGCTATGACAG 180
 QY 808 TGAAGCTGAGACCGAGAGTGAAGTGAACCGCTGTCTGTGCTGTGAGAAATGGGCTG 867
 DB 181 TGAAGCTGAGACCGAGAGTGAAGTGAACCGCTGTCTGTGCTGTGAGAAATGGGCTG 240
 QY 868 TACAGCCATGACCTGTGACGAGAAAGATCAGAAAGGGGGCCAGACCCAGACAGAGAGAG 927
 DB 241 TACAGCCATGACCTGTGACGAGAAAGATCAGAAAGGGGGCCAGACCCAGACAGAGAGAG 300
 QY 928 GATGACCAATATGTCCAGAGAGCTCCAAAAGCATGAGAAACAGCTGAAAAGACCAAGAG 987
 DB 301 GATGACCAATATGTCCAGAGAGCTCCAAAAGCATGAGAAACAGCTGAAAAGACCAAGAG 360
 QY 988 AGTGAACACCAAGAGATTAATGAAGAGCAGACAGAGTGTGATGCCAGATCTT 1047
 DB 361 AGTGAACACCAAGAGATTAATGAAGAGCAGACAGAGTGTGATGCCAGATCTT 420
 QY 1048 CTCACCTTACGCGCTGAGTTCAGTATACCAAGTGTGCTACAGTGTGCGCAATATACAG 1107
 DB 421 CTCACCTTACGCGCTGAGTTCAGTATACCAAGTGTGCTACAGTGTGCGCAATATACAG 480
 QY 1108 TATTTGCTTATATAGCAATGAGTTTATTTTGTATTTGTTTGCATTAAGATATGA 1167
 DB 481 TATTTGCTTATATAGCAATGAGTTTATTTTGTATTTGTTTGCATTAAGATATGA 540
 QY 1168 AGTGTGCTGCTAGAGAGAGAGGAGCCACAGCTTCATTTCTAGAGAGCTTTAGAGAA 1227
 DB 541 AGTGTGCTGCTAGAGAGAGAGGAGCCACAGCTTCATTTCTAGAGAGCTTTAGAGAA 600
 QY 1228 ACTGTAAATGTGTCTGTGGGCTGAGAGCTAGTAAGAAATGCTACAGATTGAAAGAG 1287
 DB 601 ACTGTAAATGTGTCTGTGGGCTGAGAGCTAGTAAGAAATGCTACAGATTGAAAGAG 660
 QY 1288 GAAACAGACCCAAATCTGAACTCTTTTGTGATTACTGTCTGTCAAGAGCTGTGAGGGA 1347
 DB 661 GAAACAGACCCAAATCTGAACTCTTTTGTGATTACTGTCTGTCAAGAGCTGTGAGGGA 720
 QY 1348 GTGACACAGATGCGACAGAGAACTTGTAGCAGAGTGTCCCGAGAGAGAGGTTTGGAGACT 1407
 DB 721 GTGACACAGATGCGACAGAGAACTTGTAGCAGAGTGTCCCGAGAGAGAGGTTTGGAGACT 780
 QY 1408 CCAAGAGAGAGAAAGCTCTCTGTCTTCCAGCTCTTTTCCATTGTGCGTCAAGATGAGACC 1467
 DB 781 CCAAGAGAGAGAAAGCTCTCTGTCTTCCAGCTCTTTTCCATTGTGCGTCAAGATGAGACC 840
 QY 1468 TCCAGATCCAGCATCTCTGTGCTCCCAATTAATCTGCTCTAGATACATAGCATCTCTCT 1527
 DB 841 TCCAGATCCAGCATCTCTGTGCTCCCAATTAATCTGCTCTAGATACATAGCATCTCTCT 900
 QY 1528 AGTTAAACCAAGTGTCTCTGTGAGCTGTGATGAGATTTCTGTGAGAGGTTACACCCAAATGATG 1587
 DB 901 AGTTAAACCAAGTGTCTCTGTGAGCTGTGATGAGATTTCTGTGAGAGGTTACACCCAAATGATG 960
 QY 1588 CAGATCTTGTATCTTGTAGGCTTGTAGGAGCCTTACACAAATTTTAAATACTTTTAA 1647
 DB 961 CAGATCTTGTATCTTGTAGGCTTGTAGGAGCCTTACACAAATTTTAAATACTTTTAA 1020

QY 1648 CCAAGAGTCTATTTCTCTGTGAAAACACTTTTTTTTGGCAAGTGTACTTATCTTCAT 1707
 DB 1021 CCAAGAGTCTATTTCTCTGTGAAAACACTTTTTTTTGGCAAGTGTACTTATCTTCAT 1079
 QY 1708 TATTTATCATTAATTAATGTTTATTTTAAATTTTATTTTCTGACAGGTATTAAGCTTTT 1767
 DB 1080 TATTTATCATTAATTAATGTTTATTTTAAATTTTATTTTCTGACAGGTATTAAGCTTTT 1139
 QY 1768 GTATATTTTTCAGTAGTGTCCCAACATTTGATAGTGAAGAGGTTGGGGTCTTCTCG 1827
 DB 1140 GTATATTTTTCAGTAGTGTCCCAACATTTGATAGTGAAGAGGTTGGGGTCTTCTCG 1199
 QY 1828 GTGACAGGGCTGAAATTAACCCAGATGCTCCACCCGACACATACATGAGAGCCATA 1887
 DB 1200 GTGACAGGGCTGAAATTAACCCAGATGCTCCACCCGACACATACATGAGAGCCATA 1259
 QY 1888 GTTGGCCCCCTAGCTTCCAGACATCTATCTGTCCAGAGAGAGAGGAGGCTTTAGAC 1947
 DB 1260 GTTGGCCCCCTAGCTTCCAGACATCTATCTGTCCAGAGAGAGAGGAGGCTTTAGAC 1319
 QY 1948 CTAGCCAGAGGAGAGAGATC-TTCATTAATAACCTTCAAGATCCAAATTAATTTGT 2006
 DB 1320 CTAAGCCAGAGGAGAGAGATCTTTCATTAATAACCTTCAAGATCCAAATTAATTTGT 1379
 QY 2007 TTTTATTTATTTCTGAGAGTGAAGCAATCAGTATTTCCAAAGATGGCGACAAGGGCAG 2066
 DB 1380 TTTTATTTATTTCTGAGAGTGAAGCAATCAGTATTTCCAAAGATGGCGACAAGGGCAG 1439
 QY 2067 CCAAGAGAGGCTT-AGGATATCCAGCCCTACCAATATGCTCAATGAGTCACT-AGGAG 2124
 DB 1440 CCAAGAGAGGCTTAAATAATTTCCAGCCCTACCAATATGCTCAATGAGTCACT-AGGAG 1499
 QY 2125 GTGAGTGGCCCTGTCTCTCTTTTCT-AGACCTCACTTCTCA- --GTGAGCTGT 2180
 DB 1500 GTGAGTGGCCCTGTCTCTCTTTTCTTTTGGAGCCCTCACTTCTCAAGTGAAGCCCTGT 1559
 QY 2181 AAGATGCACTAACCTTTGATTTGATTAAGTATTAATCTGTGCTGTGATCATTTGCTG 2240
 DB 1560 AAGATGCACTAACCTTTGATTTGATTAAGTATTAATCTGTGCTGTGATCATTTGCTG 1619
 QY 2241 CAGAGGAGAGATAGTGTCTGTGATTTTCTCTCTCTCTATAGATTAATGAATCTTG 2300
 DB 1620 CAGAGGAGAGATAGTGTCTGTGATTTTCTCTCTCTCTCTATAGATTAATGAATCTTG 1679
 QY 2301 TTACTAGACAAAGAAATGTCAATGTGCGCAAAAACAAAGTACAGATTGATCTAGCCT 2360
 DB 1680 TTACTAGACAAAGAAATGTCAATGTGCGCAAAAACAAAGTACAGATTGATCTAGCCT 1739
 QY 2361 GATGACCTTACAGAGTGTGCTATGATGAGAGTCTCATGGGTAAAGCAGAGAGAGT 2420
 DB 1740 GATGACCTTACAGAGTGTGCTATGATGAGAGTCTCATGGGTAAAGCAGAGAGAGT 1799
 QY 2421 GAAAAGAGAACCAACCCACTCTGTCTTCAATATTTGATTTCACTGCTTAACTCCGCTG 2480
 DB 1800 GAAAAGAGAACCAACCCACTCTGTCTTCAATATTTGATTTCACTGCTTAACTCCGCTG 1859
 QY 2481 AAATAGAAAGCATTTCCCTTGAAGATGAGATTAAGAAAGTTTCAAGTTCAACAGGGGGA 2540
 DB 1860 AAATAGAAAGCATTTCCCTTGAAGATGAGATTAAGAAAGTTTCAAGTTCAACAGGGGGA 1919
 QY 2541 AGAAATGAGATTTAATCTTAAACTGTGACTTGTGGAGAGTCAATCAATTTACAGTTAGT 2600
 DB 1920 AGAAATGAGATTTAATCTTAAACTGTGACTTGTGGAGAGTCAATCAATTTACAGTTAGT 1979
 QY 2601 CCTGTGTCTTGTGACTTGTGATTTAATACCCCACTCACTACCTGTTTCAAGTCAAT 2660
 DB 1980 CCTGTGTCTTGTGACTTGTGATTTAATACCCCACTCACTACCTGTTTCAAGTCAAT 2039
 QY 2661 TGAATATCCAAAGATTAATCTTGAATTAAGATCTCATTTGAGAGAGAGATTAAGA 2720
 DB 2040 TGAATATCCAAAGATTAATCTTGAATTAAGATCTCATTTGAGAGAGAGATTAAGA 2099

```

>Y 2721 CCATCAGAGGAAATATTATTAGTTCTTAATGACAGGCAACTGTGAGAACTGTGTGCCC 2780
>B 2100 CCATCAGAGGAAATATTATTAGTTCTTAATGACAGGCAACTGTGAGAACTGTGTGCCC 2159

>Y 2781 AAAAATAGAAATTCCTTCTAGT 2801
>B 2160 AAAAATAGAAATTCCTTCTAGT 2180

RESULT 7
LOCUS X410641
DEFINITION Sequence 3288 from Patent WO229103.
ACCESSION AX410641
VERSION AX410641.1 GI:21443346
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
  Alvares, C., Horne, D., Peres-da-Silva, S. and Vockley, J. G.
  Gene expression profiles in liver cancer
  Patent: WO 0229103-A 3288 11-APR-2002;
  GENE LOGIC INC (US)
FEATURES
  source          1..1987
                  /organism="Homo sapiens"
                  /mol_type="unassigned DNA"
                  /db_xref="taxon:9606"
                  /note="EMBL/GenBank Accession No. U06863"

ORIGIN
Query Match          52.5%; Score 1940; DB 6; Length 1987;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1954; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

>Y 12 CTCACCTCCGCTTACAGTCTGCTGCGCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCTG 71
>B 21 CTCACCTCCGCTTACAGTCTGCTGCGCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCTG 80

>Y 72 ACCAGACCAAGTGTGAACGCTGGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 131
>B 81 ACCAGACCAAGTGTGAACGCTGGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 140

>Y 132 GGGTCCCGCGCGGAGAGAGCTAAGAGCAAAATCCAGATCTGTGCCAATGTGTTTGTG 191
>B 141 GGGTCCCGCGCGGAGAGAGCTAAGAGCAAAATCCAGATCTGTGCCAATGTGTTTGTG 200

>Y 192 GAGCGCGCGGGAATGTGCAGTACAGAGAAAGGGNACCCACCTGCTCTGCAATTGAGC 251
>B 201 GAGCGCGCGGGAATGTGCAGTACAGAGAAAGGGNACCCACCTGCTCTGCAATTGAGC 260

>Y 252 AATGCAAACTCTCAAGAGGCGCTGTGTGGCAGTAAATGGCAAGACTACCTCAACCACT 311
>B 261 AATGCAAACTCTCAAGAGGCGCTGTGTGGCAGTAAATGGCAAGACTACCTCAACCACT 320

>Y 312 GTGAATGTGATGTGAGATGCTGCGCTCTACTGATCCCAAAATCCAGTTGATTCAGTGGAC 371
>B 321 GTGAATGTGATGTGAGATGCTGCGCTCTACTGATCCCAAAATCCAGTTGATTCAGTGGAC 380

>Y 372 ACTGCAAGAGAGAAATCCGTAAGTCCATCTGCGCAGCCAGTTGTTGCTATCAGTCCA 431
>B 381 ACTGCAAGAGAGAAATCCGTAAGTCCATCTGCGCAGCCAGTTGTTGCTATCAGTCCA 440

>Y 432 ACCGTGATGAGCTCCGAGTGCATCTCATGCTGAGTGGCTGGAAGCTGAGATCATCTCAGATG 491
>B 441 ACCGTGATGAGCTCCGAGTGCATCTCATGCTGAGTGGCTGGAAGCTGAGATCATCTCAGATG 500

>Y 492 GCTGTTCTCTAAAGGAGCAACTACAGTGAATCCCTAGCAAGTATTTTAAAGAACTTTG 551
>B 501 GCTGTTCTCTAAAGGAGCAACTACAGTGAATCCCTAGCAAGTATTTTAAAGAACTTTG 560

```

```

>Y 552 ATAAATGTGATTTCTGCTGCGACTCCAGTGAATCTCTGAAGTTTGTGGACAGAAATGAAA 611
>B 561 ATAAATGTGATTTCTGCTGCGACTCCAGTGAATCTCTGAAGTTTGTGGACAGAAATGAAA 620

>Y 612 CTGCGATCAATATTACAACTATCCAGACCCAGAGAACAACTAGTTGCTTTAGGGGACTCT 671
>B 621 CTGCGATCAATATTACAACTATCCAGACCCAGAGAACAACTAGTTGCTTTAGGGGACTCT 680

>Y 672 GTGTTGATGCTCTCAATTGAATGCTGATGAAAATGCTGATGGAAAATCAGCTTCCAAAG 731
>B 681 GTGTTGATGCTCTCAATTGAATGCTGATGAAAATGCTGATGGAAAATCAGCTTCCAAAG 740

>Y 732 AGTTTCTCAAGTGGCTCAACCCATCTTTCAACCTCTCTGAGAAAGAGTGTGCCCTGAGG 791
>B 741 AGTTTCTCAAGTGGCTCAACCCATCTTTCAACCTCTCTGAGAAAGAGTGTGCCCTGAGG 800

>Y 792 ATGAAACGTATGCAAGTGGAGCTGAGACCGAGGTGAGCTGTAAACCGCTGTGTCTGSCCT 851
>B 801 ATGAAACGTATGCAAGTGGAGCTGAGACCGAGGTGAGCTGTAAACCGCTGTGTCTGSCCT 860

>Y 852 GTGAAAATTTGGTCTGTACAGCCATGACCTGTGACGAAAAGAAATCAGAAAGGGGCCCAGA 911
>B 861 GTGAAAATTTGGTCTGTACAGCCATGACCTGTGACGAAAAGAAATCAGAAAGGGGCCCAGA 920

>Y 912 CCCAGACAGAGGAGAGATGACCAAGATATGTCCAGGAGCTCCAAAAGCATCAGGAAACAG 971
>B 921 CCCAGACAGAGGAGAGATGACCAAGATATGTCCAGGAGCTCCAAAAGCATCAGGAAACAG 980

>Y 972 CTGAAAAGACCAAGAGAGTGGACCAAGAGATCTAATGAGGAGGCACAGACCAAGTGC 1031
>B 981 CTGAAAAGACCAAGAGAGTGGACCAAGAGATCTAATGAGGAGGCACAGACCAAGTGC 1040

>Y 1032 TGGATCCCGAGCATCTTCTCCACTTCAGCGCTGAGTTTCTGATATACAAAGTGTGTGTACA 1091
>B 1041 TGGATCCCGAGCATCTTCTCCACTTCAGCGCTGAGTTTCTGATATACAAAGTGTGTGTACA 1100

>Y 1092 GTGCGCAATACCAAGTATTTGCTTATATAGCAATGATGTTTATTTGTTTATTTGTTT 1151
>B 1101 GTGCGCAATACCAAGTATTTGCTTATATAGCAATGATGTTTATTTGTTTATTTGTTT 1160

>Y 1152 GCAATAAAGGATATGAAGTGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1211
>B 1161 GCAATAAAGGATATGAAGTGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1220

>Y 1212 GAGTGTCTTTAAGAGAACTGTAAATGCTGCTCTGGGCTGGAGGCTAGTAAGGAACTGC 1271
>B 1221 GAGTGTCTTTAAGAGAACTGTAAATGCTGCTCTGGGCTGGAGGCTAGTAAGGAACTGC 1280

>Y 1272 ATCAGATTGAAAGAGGAAACAGACCCAAATCTGAACTCTTTTGAAGTTTACTGCATCTGT 1331
>B 1281 ATCAGATTGAAAGAGGAAACAGACCCAAATCTGAACTCTTTTGAAGTTTACTGCATCTGT 1340

>Y 1332 CAGCAGGCTGAGGAGGTGACACGATGCCAGAGAGAACTTAGCAGGCTGTCCCGGAGG 1391
>B 1341 CAGCAGGCTGAGGAGGTGACACGATGCCAGAGAGAACTTAGCAGGCTGTCCCGGAGG 1400

>Y 1392 AGAGTTTGGGAAGCTCCAGGAGAGGAAACCTCTCTGCTTCCAGCTCTTTCCATTGCCC 1451
>B 1401 AGAGTTTGGGAAGCTCCAGGAGAGGAAACCTCTCTGCTTCCAGCTCTTTCCATTGCCC 1460

>Y 1452 GTGAGATGACAGACCTCCAGCATCCAGCATCCAGCATCCAGCATCCAGCATCCAGCAT 1511
>B 1461 GTGAGATGACAGACCTCCAGCATCCAGCATCCAGCATCCAGCATCCAGCATCCAGCAT 1520

>Y 1512 ACATAGCCACTACTGCTAGTTAAACCCAGTGTCCCTCAGACTTGGAGTTTCTGGGAGG 1571
>B 1521 ACATAGCCACTACTGCTAGTTAAACCCAGTGTCCCTCAGACTTGGAGTTTCTGGGAGG 1580

>Y 1572 GTACACCCAAATGATGAGATACCTGTATACCTTTGAGCCCTTAGCGACCTTAACCAAT 1631
>B 1581 GTACACCCAAATGATGAGATACCTGTATACCTTTGAGCCCTTAGCGACCTTAACCAAT 1640

>Y 1632 TTAATAATACCTTTTACCAAGAGTGTCTTCTCTGTAAACACACTTTTTTTTGGCAAGTT 1691

```

```

1641 1TAAATAACATTTTACCAAGTGCTATTTCTCTGTAAAACAC-TTTTTTGGCAAGTT 1699
QY 1692 GACTTATTTCTCAATTAATATCATTAATATATGTTTTTAAATATTTATTTCTTGAC 1751
Db 1700 GACTTATTTCTCAATTAATATCATTAATATGTTTTTAAATATTTATTTCTTGAC 1759
QY 1752 TAGGTATTAAGCTTTTGAATATTTTTCAGTAGTCCACACATTCATAGTGAAGGAG 1811
Db 1760 TAGGTATTAAGCTTTTGAATATTTTTCAGTAGTCCACACATTCATAGTGAAGGAG 1819
QY 1812 TTTGGGTTCTTCCTGGTGCGAGGGCTGAATAACCCAGATGCTCCACCTGCCACATA 1871
Db 1820 TTTGGGTTCTTCCTGGTGCGAGGGCTGAATAACCCAGATGCTCCACCTGCCACATA 1879
QY 1872 CTAGATGAGCCCATAGTTGGCCCCCTTAGCTTCAGCAGTCCACTATCTGCCAGAGGAG 1931
Db 1880 CTAGATGAGCCCATAGTTGGCCCCCTTAGCTTCAGCAGTCCACTATCTGCCAGAGGAG 1939
QY 1932 CAAGGTGCTTGAACCTTAAGCCAGGGGGAAGAAGCATCTT 1971
Db 1940 CAAGGTGCTTGAACCGAAGCAGGGGAAGAAGCATCTT 1979

RESULT 8
LOCUS HSU06863 1987 bp mRNA linear PRI 11-MAY-1995
DEFINITION Human follistatin-related protein precursor mRNA, complete cds.
ACCESSION U06863
VERSION U06863.1 GI:536897
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Zwijnsen,A., Blokkx,H., Van Arnhem,W., Willems,J., Franssen,L.,
Devos,K., Raymakers,J., Van de Voorde,A. and Slegers,H.
Characterization of a rat C6 glioma-secreted follistatin-related
protein (FRP). Cloning and sequence of the human homologue
Eur. J. Biochem. 225 (3), 937-946 (1994)
95045570
PUBMED 7957230
AUTHORS Zwijnsen,A.
TITLE Direct Submission
JOURNAL Submitted (17-FEB-1994) An Zwijnsen, Biochemistry, University of
Antwerp UTA, Universiteit Antwerpen, Wilrijk, Antwerp, B-2610, Belgium
FEATURES
source
1..1987
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_lines="Hs683"
/tissue_type="glioma"
/tissue_type="brain"
1..91
5'UTR 92..1018
CDS 92..1018
/product="follistatin-related protein precursor"
/protein_id="AAA66062.1"
/db_xref="GI:536898"
/translation="MWKELALALVAVWVRAEELRSKSKICANVFCAGRECAV
TEKGEPTCLCEQKPHRPGVCSNGKTYLNHCLHRDACLTSKIDVDYDGHCKBK
SVSPASPVVCYQNDLEIRRIITQWLEAIIIPGWFSKGSNYSLEIDKFKNFNDNG
SRLDSEFLKVFQNETAINTIYPQDNKKLRLGLCVDALEISDENADWKLSFQEF
LKCLNPSNPFPEKKCALEDEYADGAETVDCNRCVACGNCWVCTAMTCGRKQKGAQ
TQTEEMTRYVQELQKHQETAETKRVSTKEI"
92..151
sig_peptide
mat_peptide 152..1015
/product="follistatin-related protein"
1019..1987
3'UTR
ORIGIN

```

```

Query Match 52.5%; Score 1940; DB 9; Length 1987;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1954; Conservative 0; Mismatches 5; Indels 1; Gaps 1;
QY 12 CTCACACCTCCGCTTACAGCTCGCTGCCCGCGCTCGCTGCCCGCGCGCCAGGAGACCTGG 71
Db 21 CTCACACCTCCGCTTACAGCTCGCTGCCCGCGCTCGCTGCCCGCGCGCCAGGAGACCTGG 80
QY 72 ACCACACACAGATGTGGAAACGCTCGCTCGCGCTCGCTGCCCGCTCGCTGCCCGCTCGCT 131
Db 81 ACCACACACAGATGTGGAAACGCTCGCTCGCGCTCGCTGCCCGCTCGCTGCCCGCTCGCT 140
QY 132 GGGTCGGCGCCGAGAGAGCTAAGAGCAAAATCAAGATCTGTGCCAATGTGTTTGTG 191
Db 141 GGGTCGGCGCCGAGAGAGCTAAGAGCAAAATCAAGATCTGTGCCAATGTGTTTGTG 200
QY 192 GAGCGCGCGCGGGAATGTGCAGTCACAGAGAAGGGGAACCCACCTGTCTCTGCAATTGAGC 251
Db 201 GAGCGCGCGCGGGAATGTGCAGTCACAGAGAAGGGGAACCCACCTGTCTCTGCAATTGAGC 260
QY 252 AATGCAACCTCACAAGAGGCTGTGTGGCAGTAATGGCAGACCTACCTCAACCACT 311
Db 261 AATGCAACCTCACAAGAGGCTGTGTGGCAGTAATGGCAGACCTACCTCAACCACT 320
QY 312 GTGAATGTCATCGAGATGCTGCTCACTGGATCCAAAATCCAGGTTGATTACGATGGAC 371
Db 321 GTGAATGTCATCGAGATGCTGCTCACTGGATCCAAAATCCAGGTTGATTACGATGGAC 380
QY 372 ACTGCAAGAGAAGAAATCCGTAAGTCCATCTGCCAGCCAGTGTGTTGCTATCAGTCA 431
Db 381 ACTGCAAGAGAAGAAATCCGTAAGTCCATCTGCCAGCCAGTGTGTTGCTATCAGTCA 440
QY 432 ACCGTGATGAGTCCGAGCTGCATCATCTGAGTCCAGGCTCGAAGCTCAGATCATCTCAGATG 491
Db 441 ACCGTGATGAGTCCGAGCTGCATCATCTGAGTCCAGGCTCGAAGCTCAGATCATCTCAGATG 500
QY 492 GCTGGTCTCTAAAGCGACAACTACAGTGAATCCTAGACAAGATTTTAAAGACTTTG 551
Db 501 GCTGGTCTCTAAAGCGACAACTACAGTGAATCCTAGACAAGATTTTAAAGACTTTG 560
QY 552 ATATGGTGATTTCTCGCTGGACTCCAGTGAATCTCTGAAGTTTGTGGAACAGATGAAA 611
Db 561 ATATGGTGATTTCTCGCTGGACTCCAGTGAATCTCTGAAGTTTGTGGAACAGATGAAA 620
QY 612 CTGCCATCAATATTACAGTATCCAGACCCAGGAGAACCAAGTTGCTTAGGGAGCTCT 671
Db 621 CTGCCATCAATATTACAGTATCCAGACCCAGGAGAACCAAGTTGCTTAGGGAGCTCT 680
QY 672 GTGTGATGCTCTCATTTGAATCTGTGATGAAATGCTGATGAAATCTCAGCTTCCAG 731
Db 681 GTGTGATGCTCTCATTTGAATCTGTGATGAAATGCTGATGAAATCTCAGCTTCCAG 740
QY 732 AGTTTCTCAAGTGCTCAACCCATCTTTCAACCCCTCTGAGAGAAGTGTGCCCTGGAGG 791
Db 741 AGTTTCTCAAGTGCTCAACCCATCTTTCAACCCCTCTGAGAGAAGTGTGCCCTGGAGG 800
QY 792 ATGAACAGTATCCAGATGAGCTGAGACCGAGGTGAGCTGTAAACCGCTGTGTCTGTGCT 851
Db 801 ATGAACAGTATCCAGATGAGCTGAGACCGAGGTGAGCTGTAAACCGCTGTGTCTGTGCT 860
QY 852 GTGGAATTTGGTCTGTACAGCCATGACCTGTGACGGAAGAAATCAGAAGGGGGGCCAGA 911
Db 861 GTGGAATTTGGTCTGTACAGCCATGACCTGTGACGGAAGAAATCAGAAGGGGGGCCAGA 920
QY 912 CCAGACAGAGAGAGATGACAGATATGTCCAGAGCTCCAAAGATCAGGAAACAG 971
Db 921 CCAGACAGAGAGAGATGACAGATATGTCCAGAGCTCCAAAGATCAGGAAACAG 980
QY 972 CTGAAAGACCAAGAGAGTGAGCAACCAAGAGATCTAATGAGAGGCGCAGACAGTGTCT 1031
Db 981 CTGAAAGACCAAGAGAGTGAGCAACCAAGAGATCTAATGAGAGGCGCAGACAGTGTCT 1040

```

1032	2y	TGGATCCAGCATCTTCTTCACCTTCAGCGCTGAGTTCAGTATACACAAGTGTCTGCTACA	1091
1041	2b	TGGATCCAGCATCTTCTTCACCTTCAGCGCTGAGTTCAGTATACACAAGTGTCTGCTACA	1100
1092	2y	GTCCGCAAAATCACCAAGTATTTGCTTATATAGCAATGAGTTTATTTTGTATTTTGTGTTT	1151
1101	2b	GTCCGCAAAATCACCAAGTATTTGCTTATATAGCAATGAGTTTATTTTGTATTTTGTGTTT	1160
1152	2y	GCAATTAAGGATATGAAGTGTGGCTGGCTAGGAAGGNAAGGCCACAGCCTTCATTCTCTAG	1211
1161	2b	GCAATTAAGGATATGAAGTGTGGCTGGCTAGGAAGGNAAGGCCACAGCCTTCATTCTCTAG	1220
1212	2y	GAGTGCCTTTAAGAGAAACTCTGTAATGGTGTCTCTGGGCTGGAGGCTAGTAAGGAAACTGC	1271
1221	2b	GAGTGCCTTTAAGAGAAACTCTGTAATGGTGTCTCTGGGCTGGAGGCTAGTAAGGAAACTGC	1280
1272	2y	ATCAGGNTGAAAGAGGAAACAGACCCCAAAATCTGAACTCTCTTTTGTAGTTTACTGCATCTGT	1331
1281	2b	ATCAGGNTGAAAGAGGAAACAGACCCCAAAATCTGAACTCTCTTTTGTAGTTTACTGCATCTGT	1340
1332	2y	CAGCAGGCTGCAGGGAGTGCACACGATGCCAGAGAGAACTTAGCAGGGTGTCCCGGAGG	1391
1341	2b	CAGCAGGCTGCAGGGAGTGCACACGATGCCAGAGAGAACTTAGCAGGGTGTCCCGGAGG	1400
1392	2y	AGAGGTTTGGGAGCTCCAGGGAGAGAAACGCTCTCTGCTTCCAGCCTCTTCCCATGTCC	1451
1401	2b	AGAGGTTTGGGAGCTCCAGGGAGAGAAACGCTCTCTGCTTCCAGCCTCTTCCCATGTCC	1460
1452	2y	GTGAGCATGACAGACCTCCAGCATCCACGCATCTCTTGGTCCCAATAAAGTCCCTTAGAT	1511
1461	2b	GTGAGCATGACAGACCTCCAGCATCCACGCATCTCTTGGTCCCAATAAAGTCCCTTAGAT	1520
1512	2y	ACATAGCCATCTGCTAGTTTAAACCCAGTGTCCCTCAGACTTGGATGGAGTTTCTGGGAGG	1571
1521	2b	ACATAGCCATCTGCTAGTTTAAACCCAGTGTCCCTCAGACTTGGATGGAGTTTCTGGGAGG	1580
1572	2y	GTACACCCAAATGATGCAGATACTTGTTATCTTTGAGCCCTTAGCGACCTAACCCAAAT	1631
1581	2b	GTACACCCAAATGATGCAGATACTTGTTATCTTTGAGCCCTTAGCGACCTAACCCAAAT	1640
1632	2y	TTAAATAATCTTTTACCAAGGTGCTATTTCTCTGTAAACACACTTTTTTTTGGCAAGTT	1691
1641	2b	TTAAATAATCTTTTACCAAGGTGCTATTTCTCTGTAAACACACTTTTTTTTGGCAAGTT	1699
1692	2y	GACTTTATCTTCAATPATATCAATATATATTTGTTTTTAAATATTTTATTTCTTTGAC	1751
1700	2b	GACTTTATCTTCAATPATATCAATATATATTTGTTTTTAAATATTTTATTTCTTTGAC	1759
1752	2y	TAGGTATTAAGCTTTTCTGTAATATTTTTCAGTAGTCCACACCTTCATAGTGGAGGAG	1811
1760	2b	TAGGTATTAAGCTTTTCTGTAATATTTTTCAGTAGTCCACACCTTCATAGTGGAGGAG	1819
1812	2y	TTTGGGTTCTTCTGTGTGCGAGGGCTGAAATAACCCAGATGTCCCAACCTTGCACATA	1871
1820	2b	TTTGGGTTCTTCTGTGTGCGAGGGCTGAAATAACCCAGATGTCCCAACCTTGCACATA	1879
1872	2y	CTAGATCAGCCCATAGTTGTGCCCCCTTAGCTTCCAGCAGTCCACTATCTGCCAGAGGAG	1931
1880	2b	CTAGATCAGCCCATAGTTGTGCCCCCTTAGCTTCCAGCAGTCCACTATCTGCCAGAGGAG	1939
1932	Qy	CAAGGGTGCCTTAGACCTTAAGCCAGGGGGAAGAAGCATCTT	1971
1940	2b	CAAGGGTGCCTTAGACCCGAAGCCAGGGGGAAGAAGCATCTT	1979
RESULT 9			
AX577955	LOCUS	AX577955	2538 bp DNA linear PAT 08-JAN-2003
DEFINITION	Sequence 77 from Patent WO02081745.		
ACCESSION	AX577955		
VERSION	AX577955.1		
KEYWORDS	GI:27647163		
SOURCE	Homo sapiens (human)		

ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 Garcia,T., roman Roman,S., Baron,R., Call,K., Theilhaber,J., Connolly,T., Jackson,A., Bushnell,S.B. and Rawadi,G.
TITLE	Genes involved in osteogenesis, and methods of use
JOURNAL	Patent: WO 02081745-A 77 17-OCT-2002;
FEATURES	Avantis Pharma S.A. (FR)
source	Location/Qualifiers
ORIGIN	1. 2538 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606"
Query Match	46.8%; Score 1729.8; DB 6; Length 2538;
Best Local Similarity	99.8%; Pred. No.0;
Matches 1742;	Conservative 0; Mismatches 2; Indels 1; Gaps 1;
QY	12 CTCACACTCCGTTACAGCTCGCTCGCCGCTCTGTCGCCCGCGCCCCCAGAGACCTGG 71
Db	6 CTCACACTCCGTTACAGCTCGCTCGCCGCTCTGTCGCCCGCGCCCCCAGAGACCTGG 65
QY	72 ACCAGACACGANGTGGAAACGCTGCTCGCGCTCGCGCTCGCGTGTGTGCGGTGCGCT 131
Db	66 ACCAGACACGANGTGGAAACGCTGCTCGCGCTCGCGCTCGCGTGTGTGCGGTGCGCT 125
QY	132 GGTTCGCGCGCGAGGAGAGCTAAGAGGAAATCCAAAGATCTGTGCCAATGTGTTTGTG 191
Db	126 GGTTCGCGCGCGAGGAGAGCTAAGAGGAAATCCAAAGATCTGTGCCAATGTGTTTGTG 185
QY	192 GAGCGCGCGGGAAATGTCAGTCCAGAGAAAGGGAAACCCACCTGTCTGTCAATGAGC 251
Db	186 GAGCGCGCGGGAAATGTCAGTCCAGAGAAAGGGAAACCCACCTGTCTGTCAATGAGC 245
QY	252 AATGCAAACTCTCAAGAGCGCTGTGTGTGGCAGTAATGGCAAGACCTACCTCAACCACT 311
Db	246 AATGCAAACTCTCAAGAGCGCTGTGTGTGGCAGTAATGGCAAGACCTACCTCAACCACT 305
QY	312 GTGAACCTGCATCGAGATGCTCGCTCACCTGGATCCAAATCCAGGTTCATTCAGATGAC 371
Db	306 GTGAACCTGCATCGAGATGCTCGCTCACCTGGATCCAAATCCAGGTTCATTCAGATGAC 365
QY	372 ACTGCAAGAGAGAAATCGTGAATGTCATCTGCAGCGCCAGTTGTTGCTATCAGTCCA 431
Db	366 ACTGCAAGAGAGAAATCGTGAATGTCATCTGCAGCGCCAGTTGTTGCTATCAGTCCA 425
QY	432 ACCGTGATGAGCTCCGACGTGCGATCATCCAGTGGCTGGAGCTGAGATCATCTCAGATG 491
Db	426 ACCGTGATGAGCTCCGACGTGCGATCATCCAGTGGCTGGAGCTGAGATCATCTCAGATG 485
QY	492 GCTGGTCTCTAAAGGAGCAACTACAGTGAATCTCTAGCAAGTATTTTAAAGAACTTTG 551
Db	486 GCTGGTCTCTAAAGGAGCAACTACAGTGAATCTCTAGCAAGTATTTTAAAGAACTTTG 545
QY	552 ATAATGTTGATCTTCGCTGGACTCCAGTGAATCTCTGAAGTTTGTGGAAACAGAAATGAAA 611
Db	546 ATAATGTTGATCTTCGCTGGACTCCAGTGAATCTCTGAAGTTTGTGGAAACAGAAATGAAA 605
QY	612 CTGCCATCAATATTAACAATATCCAGACAGAGAGAACAAAGTGTGCTTAGGGGACTCT 671
Db	606 CTGCCATCAATATTAACAATATCCAGACAGAGAGAACAAAGTGTGCTTAGGGGACTCT 665
QY	672 GTGTTGATGCTCTCATGAACTGCTGTGATGAAATGCTGATTGGAAACTCAGTCTCCAAG 731
Db	666 GTGTTGATGCTCTCATGAACTGCTGTGATGAAATGCTGATTGGAAACTCAGTCTCCAAG 725
QY	732 AGTTTCTCAAGTGCTCAACCCATCTTTCAACCTCTCTGAGAGAGAGTGCCTCTGGAG 791
Db	726 AGTTTCTCAAGTGCTCAACCCATCTTTCAACCTCTCTGAGAGAGAGTGCCTCTGGAG 785
QY	792 ATGAAACGTATGCAGATGGAGCTGACACCGAGGTGGACTGTAAACCGCTGTCTGTGCT 851

Db 786 ATGAAACGTTATGCAATGGAGCTGAGACCGAGGTGGAATGTAACCGCTGTGTCTGTCCT 845
 Qy 852 GTGGAATTTGGGTCTGTACACCCATGCTGTGTGACGGAAGAAATCAGAAAGGGGGCCAG 911
 Db 846 GTGGAATTTGGGTCTGTACACCCATGCTGTGTGACGGAAGAAATCAGAAAGGGGGCCAG 905
 Qy 912 CCCAGACAGAGGAGGATGACCAAGATATGTCAGAGGAGCTCCAAAGCATCAGAAACAG 971
 Db 906 CCCAGACAGAGGAGGATGACCAAGATATGTCAGAGGAGCTCCAAAGCATCAGAAACAG 965
 Qy 972 CTGAAAGACCAAGAGAGTGAGCACCACCAAGAGATCTAATGAGGAGGACACAGACAGTGC 1031
 Db 966 CTGAAAGACCAAGAGAGTGAGCACCACCAAGAGATCTAATGAGGAGGACACAGACAGTGC 1025
 Qy 1032 TGGATCCAGCATCTTCCCACTTCAGCGCTGAGTTCAGTATACACAGTGTCTGTCTACA 1091
 Db 1026 TGGATCCAGCATCTTCCCACTTCAGCGCTGAGTTCAGTATACACAGTGTCTGTCTACA 1085
 Qy 1092 GTCGCCAAATCACCAGTATTTGCTTATATAGCAATGATTTATTTGTTTATTTGTTT 1151
 Db 1086 GTCGCCAAATCACCAGTATTTGCTTATATAGCAATGATTTATTTGTTTATTTGTTT 1145
 Qy 1152 GCATTAAGGATATCAAGTGGCTGGCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1211
 Db 1146 GCATTAAGGATATCAAGTGGCTGGCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1205
 Qy 1212 GAGTCTTTTAAAGAGAACTGTAATATGCTCTGCGGCTGAGGAGTGTAGTAAAGAAATCG 1271
 Db 1206 GAGTCTTTTAAAGAGAACTGTAATATGCTCTGCGGCTGAGGAGTGTAGTAAAGAAATCG 1265
 Qy 1272 ATCAGGATGAAAGAGGAAACAGACCCAAATCTGAACCTCTTTTACGTTTACTGATCTGT 1331
 Db 1266 ATCAGGATGAAAGAGGAAACAGACCCAAATCTGAACCTCTTTTACGTTTACTGATCTGT 1325
 Qy 1332 CAGCAGGCTGACGAGGAGTGCACACCATGCCAGAGAGAACTTAGCAGGAGTGTCCCGGAGG 1391
 Db 1326 CAGCAGGCTGACGAGGAGTGCACACCATGCCAGAGAGAACTTAGCAGGAGTGTCCCGGAGG 1385
 Qy 1392 AGAGTTTGGGAGCTCCAGGAGAGGAAAGCTCTCTGCTTCCAGGCTCTTTCCATTTGCC 1451
 Db 1386 AGAGTTTGGGAGCTCCAGGAGAGGAAAGCTCTCTGCTTCCAGGCTCTTTCCATTTGCC 1445
 Qy 1452 GTCAGCATGACAGACTCCAGCATCCAGCATCTCTGCTGCCAATTAAGTCTCTAGAT 1511
 Db 1446 GTCAGCATGACAGACTCCAGCATCCAGCATCTCTGCTGCCAATTAAGTCTCTAGAT 1505
 Qy 1512 ACATAGCCATCTGTAGTTAACCCAGTGTCCCTCAGACTTGGATGGAGTTTCTGGAGG 1571
 Db 1506 ACATAGCCATCTGTAGTTAACCCAGTGTCCCTCAGACTTGGATGGAGTTTCTGGAGG 1565
 Qy 1572 GTACCCCAATGATGACGATCTGTATCTTGTAGCCCTTAGCGACCTAACCAAT 1631
 Db 1566 GTACCCCAATGATGACGATCTGTATCTTGTAGCCCTTAGCGACCTAACCAAT 1625
 Qy 1632 TTAATAATCTTTTACCAAGAGTGTATTTCTCTGTAAAAAC-TTTTTTTTGGCAAGT 1690
 Db 1626 TTAATAATCTTTTACCAAGAGTGTATTTCTCTGTAAAAAC-TTTTTTTTGGCAAGT 1685
 Qy 1691 TCAGTTTATCTCAATTTATCATTTATTTATTTATTTATTTATTTATTTATTTTCTTGA 1750
 Db 1686 TCAGTTTATCTCAATTTATCATTTATTTATTTATTTATTTATTTATTTATTTTCTTGA 1745
 Qy 1751 CTAGG 1755
 Db 1746 CTAGG 1750

RESULT 10
 D89937
 LOCUS 2538 bp mRNA linear PRI 23-JUN-1999
 DEFINITION Homo sapiens mRNA for follistatin-related protein (FRP), complete cds.

D89937
 VERSION 1 GI:3184392
 KEYWORDS follistatin-related protein (FRP).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (sites)
 Tanaka,M., Ozaki,S., Osakada,F., Mori,K., Okubo,M. and Nakao,K.
 Cloning of follistatin-related protein as a novel autoantigen in systemic rheumatic diseases
 Int. Immunol. 10 (9), 1305-1314 (1998)
 JOURNAL 99000396
 MEDLINE 9786430
 PUBMED 9786430
 REFERENCE 2 (bases 1 to 2538)
 Tanaka,M.
 Direct Submission
 Submitted (18-DEC-1996) Masao Tanaka, Kyoto University Graduate School of Medicine, Department of Medicine and Clinical Science, 54 Shogoin-Kawahara-cho, Sakyo-ku, Kyoto, Kyoto 606-8507, Japan (E-mail:masaanaokuhp.kyoto-u.ac.jp, Tel:+81-75-751-3181, Fax:+81-75-771-9452)
 FEATURES
 Location/Qualifiers
 1..2538
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /cell_line="bulk"
 /cell_type="synovial cells"
 /tissue_type="synovium"
 /clone_lib="lambda phage"
 /dev_stage="matured"
 77..1003
 /codon_start=1
 /product="follistatin-related protein (FRP)"
 /protein_id="BAA28707.1"
 /db_xref="GI:3184393"
 /translation="MKKWLALALAVAVWRAEELRSKSKICANYFCGAGRECAV
 TEKPTCLCTBOCKPHKPVGSGKTYLHNLHRLDACLTSKIQVYDGHCKEKK
 SVSPASPVVCQSNRDLRRRIQWLEAEIIPDGMFKSGNSYSLIDKYFKNPNQD
 SRLDSBELKVEQNETAINITYDQENKLLGLCYDALIELSDENADWKLSPQEF
 LKCLNPSFPPEKKALDEITYADGAETEVCNRCVCAGCNWVCTAMTCDGNQKGAQ
 TQTEEMTRYVQELQHQETAEKTKRVSTKEI"
 77..136
 sig_peptide
 misc_feature 161..376
 /note="follistatin (FS) domain"
 377..793
 misc_feature
 /note="extracellular Ca2+ binding (EC) domain"
 692..730
 misc_feature
 /note="EF hand"
 ORIGIN
 Query Match 46.8%; Score 1729.6; DB 9; Length 2538;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1742; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
 Qy 12 CTCGCCACCTCCGTTACAGCTGCTGCCCGCGCTCTGCCCCCGCCCCCAGAGAGCCTGG 71
 Db 6 CTCGCCACCTCCGTTACAGCTGCTGCCCGCGCTCTGCCCCCGCCCCCAGAGAGCCTGG 65
 Qy 72 ACCAGACCAAGATGTGCAAAACGCTGGCTCGCGCTCGCGCTCGCGCTGTGTGGCGGTGCGCT 131
 Db 66 ACCAGACCAAGATGTGCAAAACGCTGGCTCGCGCTCGCGCTGTGTGGCGGTGCGCT 125
 Qy 132 GGGTCCGCGCCGAGGAGAGCTAAGAGCAAAATCCAGATCTGTGCCAATGTGTTTGTG 191
 Db 126 GGGTCCGCGCCGAGGAGAGCTAAGAGCAAAATCCAGATCTGTGCCAATGTGTTTGTG 185
 Qy 192 GAGCCGCGCGGGAATGTGAGTCACAGAGAAAGGGGAAACCCACCTGTCTCTGCAATTGAGC 251
 Db 186 GAGCCGCGCGGGAATGTGAGTCACAGAGAAAGGGGAAACCCACCTGTCTCTGCAATTGAGC 245
 Qy 252 AATGCAAAACCTCAAGAGGCGCTGTGTGTGGCAGTAAATGGCAAGACCTACCTCAACCACT 311

|||||
246 AATGCAAACTCACAAGAGGCGCTGTGTGGAGTAATGCGAAGACCTACCTCAACCACT 305
|||||
312 GTGAATGTCATGAGATGCTGCTCACTGGATCCAAATCAGAGTGTGATTACGATGGAC 371
|||||
306 GTGAATGTCATGAGATGCTGCTCACTGGATCCAAATCAGAGTGTGATTACGATGGAC 365
|||||
372 ACTGCAAGAGAGAAATCCGTAAGTCATCTGCGAGCCAGTGTGTTGCTATCAGTCCA 431
|||||
366 ACTGCAAGAGAGAAATCCGTAAGTCATCTGCGAGCCAGTGTGTTGCTATCAGTCCA 425
|||||
432 ACCGTGATGAGTCCGAGTCCGATCATCCAGTGGCTGGAAGCTGAGATCATCTCAGATG 491
|||||
426 ACCGTGATGAGTCCGAGTCCGATCATCCAGTGGCTGGAAGCTGAGATCATCTCAGATG 485
|||||
492 GCTGGTCTCTAAAGGAGCAATACAGTGAATCCTAGCAAGTATTTTAAAGACTTTG 551
|||||
486 GCTGGTCTCTAAAGGAGCAATACAGTGAATCCTAGCAAGTATTTTAAAGACTTTG 545
|||||
552 ATAAATGGTATTCCTCGCTGAGTCCAGTGAATTCCTGGAAGTGTGGAACAGAGTGA 611
|||||
546 ATAAATGGTATTCCTCGCTGAGTCCAGTGAATTCCTGGAAGTGTGGAACAGAGTGA 605
|||||
612 CTGCCATCAATATTACAGTATCCAGAGCAGAGAACCAAGTGTGTTAGGGACTCT 671
|||||
606 CTGCCATCAATATTACAGTATCCAGAGCAGAGAACCAAGTGTGTTAGGGACTCT 665
|||||
672 GTGTTGATGCTCTCAATGAGTCTGATGAAATGCTGATGGAATCTGGAATCTCAGTCC 731
|||||
666 GTGTTGATGCTCTCAATGAGTCTGATGAAATGCTGATGGAATCTGGAATCTCAGTCC 725
|||||
732 AGTTTCTCAAGTGCCTCAACCCATCTTCAACCCCTCTGAGAGAAAGTGTGCTCCGAGG 791
|||||
726 AGTTTCTCAAGTGCCTCAACCCATCTTCAACCCCTCTGAGAGAAAGTGTGCTCCGAGG 785
|||||
792 ATGAAGAGTATGAGATGAGTGTGAGCCGAGTGTGATGTAACCGCTGTGCTGCTCT 851
|||||
786 ATGAAGAGTATGAGATGAGTGTGAGCCGAGTGTGATGTAACCGCTGTGCTGCTCT 845
|||||
852 GTGGAATTCGGTCTGTACAGCCTACCTGTGACCGAAGATCAGAGGGGGCCAG 911
|||||
846 GTGGAATTCGGTCTGTACAGCCTACCTGTGACCGAAGATCAGAGGGGGCCAG 905
|||||
912 CCAGAGCAGAGGAGATGACAGATATGTCAGAGCTCCAAAGATCAGAGAAACAG 971
|||||
906 CCAGAGCAGAGGAGATGACAGATATGTCAGAGCTCCAAAGATCAGAGAAACAG 965
|||||
972 CTGAABAGCCAGAGTGTGAGCAGCAACAGATCTAATGAGGAGGACAGACCAAGTGC 1031
|||||
966 CTGAABAGCCAGAGTGTGAGCAGCAACAGATCTAATGAGGAGGACAGACCAAGTGC 1025
|||||
1032 TGGATCCAGCATCTTCCACTTCAGCGCTGAGTTCAGTATACAAAGTGTGCTACA 1091
|||||
1026 TGGATCCAGCATCTTCCACTTCAGCGCTGAGTTCAGTATACAAAGTGTGCTACA 1085
|||||
1092 GTGCCCAATCAGAGTATGCTTATAGCAATGAGTATTTTATTTTATTTTATTTT 1151
|||||
1086 GTGCCCAATCAGAGTATGCTTATAGCAATGAGTATTTTATTTTATTTTATTTT 1145
|||||
1152 GCATAAAGATATGAAGTGGCTGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1211
|||||
1146 GCATAAAGATATGAAGTGGCTGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1205
|||||
1212 GAGTGCTTTAAGAGAACTGTAATGCTGTGAGGCTGAGGCTGAGGCTGAGGCTGAG 1271
|||||
1206 GAGTGCTTTAAGAGAACTGTAATGCTGTGAGGCTGAGGCTGAGGCTGAGGCTGAG 1265
|||||
1272 ATCAGATTTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1331
|||||
1266 ATCAGATTTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1325
|||||
1332 CAGCAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1391
|||||

Db 1326 CAGCAGCTCAGGAGTGACACAGATGCCAGAGAACCTTAGCAGGGTGTCCCGAGG 1385
Qy 1392 AGAGTTTGGAAAGCTCCACGAGAGGAAACGCTCTCTGCTCCAGGCTCTTTCCATTGCC 1451
Db 1386 AGAGTTTGGAAAGCTCCACGAGAGGAAACGCTCTCTGCTCCAGGCTCTTTCCATTGCC 1445
Qy 1452 GTACAGATGACAGACCTCCAGCATCCAGCATCTCTGGTCCCAATAACCTGCTTAGAT 1511
Db 1446 GTACAGATGACAGACCTCCAGCATCCAGCATCTCTGGTCCCAATAACCTGCTTAGAT 1505
Qy 1512 ACATAGCCATACCTGCTAGTTAAACCCAGTGTCCCTCAGACTTGGATGGAGTTTCTGGAGG 1571
Db 1506 ACATAGCCATACCTGCTAGTTAAACCCAGTGTCCCTCAGACTTGGATGGAGTTTCTGGAGG 1565
Qy 1572 GTACACCAATGATGAGATCTGTATATCTTGTAGCCCTTAGCCCTTAGCCCTTAACCAATT 1631
Db 1566 GTACACCAATGATGAGATCTGTATATCTTGTAGCCCTTAGCCCTTAGCCCTTAACCAATT 1625
Qy 1632 TTAATAATACCTTTTACCAAGAGTGTATTTCTCTGTAAACAC-TTTTTTTGGCAAGT 1690
Db 1626 TTAATAATACCTTTTACCAAGAGTGTATTTCTCTGTAAACACACTTTTTTTTGGCAAGT 1685
Qy 1691 TGACTTTATCTTCAAT 1750
Db 1686 TGACTTTATCTTCAAT 1745
Qy 1751 CTAGG 1755
Db 1746 CTAGG 1750

RESULT 11
AK025860
LOCUS Homo sapiens cDNA: FLJ22207 fis, clone HRC01454. 1694 bp mRNA linear PRI 13-SEP-2003
DEFINITION AK025860
ACCESSION AK025860
VERSION AK025860.1 GI:10438502
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 Watanabe,K., Kumagai,A., Itakura,S., Yamazaki,M., Tashiro,H.,
Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
Nakamura,Y., Isogai,T. and Sugano,S.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 1694)
Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
Shibahara,T., Tanaka,T. and Nakamura,Y.
Direct Submission
Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure Analysis, Human
Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
Japan (E-mail: flicdn@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing; Research Association for Biotechnology; cDNA library
construction, 5'- & 3'-end one pass sequencing; Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).
Location/Qualifiers
1..1694
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HRC01454"
/cell_type="primary human renal epithelial cells"
/clone_lib="HRC"
/note="Cloning vector pME18FL3"

QY	3059	TAGAACATGGTTATCATCAAGACTACTTACCTCGCAACATTTGAACCTCCCAAGAGCGAAA	3111
Db	1020	TAGAACATGGTTATCATCAAGACTACTTACCTCGCAACATTTGAACCTCCCAAGAGCGAAA	1079
QY	3119	TCACACATTCCTCTTGAGTTCTGAGCTTCTGTGTAAATAGGCGAGCTGTGCTGTATGCGG	3178
Db	1080	TCACACATTCCTCTTGAGTTCTGAGCTTCTGTGTAAATAGGCGAGCTGTGCTGTATGCGG	1139
QY	3179	TAGAATCACATGATCTGAGGACCATTCATGGAAGCTGCTTAATATAGCTAGCTATGCGGAGT	3238
Db	1140	TAGAATCACATGATCTGAGGACCATTCATGGAAGCTGCTTAATATAGCTAGCTATGCGGAGT	1199
QY	3239	CTTCCATAAAGTTTTCATGAGGACCAACAAACAGGATTTAACTAGGTTTGGTTCCCTTCAG	3298
Db	1200	CTTCCATAAAGTTTTCATGAGGACCAACAAACAGGATTTAACTAGGTTTGGTTCCCTTCAG	1259
QY	3299	CCCTCTAAAGCATAGGCTTACCTGCAGGCTTCTTGGGCTTCTCTGTGTGTGTAGT	3358
Db	1260	CCCTCTAAAGCATAGGCTTACCTGCAGGCTTCTTGGGCTTCTCTGTGTGTGTAGT	1319
QY	3359	TTTGTAACACATATAGCATCTGTTAAGATCCAGTGTCCATGGAACCTTCCACATGCGG	3418
Db	1320	TTTGTAACACATATAGCATCTGTTAAGATCCAGTGTCCATGGAACCTTCCACATGCGG	1379
QY	3419	TGACTCTGGAATATATCAGTTTTCGTTTGGAAAGCAGGGTTCTCTGCTGCTTAACAGCCGAC	3478
Db	1380	TGACTCTGGAATATATCAGTTTTCGTTTGGAAAGCAGGGTTCTCTGCTGCTTAACAGCCGAC	1439
QY	3479	GTGACCCAGTCTCAATGTCTTCTTTCATACACCTATGTTTTTAAAGTAGTCAAACTTCAAGA	3538
Db	1440	GTGACCCAGTCTCAATGTCTTCTTTCATACACCTATGTTTTTAAAGTAGTCAAACTTCAAGA	1499
QY	3539	AACAACTTAACAAAGTTTCGTTTGCATATATGTTTGTGAACCTGTATTTGATTTAGTAG	3598
Db	1500	AACAACTTAACAAAGTTTCGTTTGCATATATGTTTGTGAACCTGTATTTGATTTAGTAG	1559
QY	3599	GCCTCTATATTCATTTAACTCTGTTTTTGAACCTCTGATTTCTCTTTTCGGATCTAT	3658
Db	1560	GCCTCTATATTCATTTAACTCTGTTTTTGAACCTCTGATTTCTCTTTTCGGATCTAT	1619
QY	3659	TGATGAATAAGAAATTTAAAGTGAAAAAATAAAAAA	3693
Db	1620	TGATGAATAAGAAATTTAAAGTGAAAAAATAAAAAA	1654
RESULT 12			
BC017413			
LOCUS	BC017413	1587 bp	linear
DEFINITION	Homo sapiens, clone IMAGE:4619382, mRNA.		
ACCESSION	BC017413		
VERSION	BC017413.1		
KEYWORDS	GI:16924272		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 1587)		
TITLE	Strausberg, R.		
JOURNAL	Direct Submission		
REMARK	Submitted (13-NOV-2001) National Institutes of Health, Mammalian		
COMMENT	Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
	NIH-MGC Project URL: http://mgc.nci.nih.gov		
	Contact: MGC help desk		
	Email: cgabs@mail.nih.gov		
	Tissue Procurement: CLONTECH		
	cDNA Library Preparation: CLONTECH Laboratories, Inc.		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)		
	DNA Sequencing by: Institute for Systems Biology		
	http://www.systemsbio.org		
	contact: amadan@systemsbiology.org		

RESULT 12	REFERENCE	REMARK
BC017413	AUTHORS	COMMENT
LOCUS	TITLE	
DEFINITION	JOURNAL	
ACCESSION		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		

[illegible]

841 AAAGGTTTTCACGAGCCCTGGACATCTGGGAGGAGGCAAGATACCAATTTGTTAA 900

3024 AAGCAAGAAACACAGTGTCTCTTCCACTAGTCATTTAGAAATGGTTATCATCCAAGACT 3083

901 AAGCAAGAAACACAGTGTCTCTTCCACTAGTCATTTAGAAATGGTTATCATCCAAGACT 960

3084 ACTCTACCTGCAACATTGAACTCCAGAGCAAAATCCCATTCCTCTTGGATTTCTGCAG 3143

961 ACTCTACCTGCAACATTGAACTCCAGAGCAAAATCCCATTCCTCTTGGATTTCTGCAG 1020

[illegible]

Altechul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.P., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Datchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Uadin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullaby,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Woxley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Faneay,J., Helton,E., Kettenan,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smilova,D.E., Scherch,A., Schein,J.E., Jones,S.J., and Marra,M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22388257
12477932

2 (bases 1 to 2851)

Strausberg,R.

Direct Submission

Submitted (01-MAY-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapbs-remail.nih.gov

Tissue procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: amg@bcm.tmc.edu

Guarantee, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAP Plate: 30 Row: a Column: 7.

FEATURES

Location/Qualifiers

1..2851

/organism="Mus musculus"

/mol_type="mRNA"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="MGC:25371 IMAGE:2647002"

/tissue_type="mammary tumor. Metallothionien-TGF alpha model. 10 month old virgin mouse. Taken by biopsy."

/clone_lib="NCI CGAP Mam1"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

1..2851

/gene="Pst1"

/note="synonym: TSC-36"

/db_xref="LocusID:14314"

/db_xref="MGI:102793"

83..1003

/codon_start=1

/product="follistatin-like"

/protein_id="AAH28921.1"

/db_xref="GI:20810033"

/db_xref="LocusID:14314"

/translation="MKRWLALSLVTIALVHGEERPSKSKI CANVFCGAGRECAVTE KGEPTCLCEQCKPKRVCVCSNGKTYLNHCELRDACTGSKIQVDYDGHCKEKSSA SPSPFVVCYQANDELRIQWLEAEIIPDFWFSKSNYSLEILDYFKSPFNGDHS LDSSSEFLKVEQNETAINITTYAQENNKILSLCVDALLIELSDENADWELSLQEFLEK

gene

CDS

CLNPSFNPEKKCALEDETYADGAETEVDNCRVCSCGHVWYCTAMTCDKKNKGQVTH
TEEEKTVQVQELQKHQTAETKKNVTKEI"

236..370

/note="KAZAL; Region: Kazal type serine protease inhibitors"

/db_xref="CDD:smart00280"

ORIGIN

Query Match 28.4%; Score 1049; DB 10; Length 2851;
Best Local Similarity 76.9%; Pred No. 2,3e-247;
Matches 1591; Conservative 0; Mismatches 390; Indels 88; Gaps 22;

QY 12 CTCACACCTCCGCTTACAGCTCGCTGCGCCGCTCTGTCGCCGCGCCGCCAGAGACCTGG 71
Db 12 CTCACACCTTCGCTTAACTCGCTGCGCCGACCCCTGCGCCAGTGTCTCCGAGTCCCGG 71

QY 72 ACCAGACACGATGTGGAAACGCTCGCTCGCGCTCGCTGCTGCTGCTGCTGCTGCTGCT 131
Db 72 ACCGAGACGATGTGGAAACGATG-----CTGCGCTCTCGCTGCTGCTGCTGCTGCTG 125

QY 132 GGTTCGCGCCGAGGAGAGCTTAAGGACAAATCCAAATCTGTGCCAAATGTTTTGTG 191
Db 126 TGGTCCACGCGGAGGAGGAACTTAGAAGCAAATCCAAATCTGCGCCAAATGTGTTTTGTG 185

QY 192 GAGCGCGCGGAATGTGCAGTCACAGAGAAGGGGACCCACCTGTCTCTGCATTGAGC 251
Db 186 GAGCTGGCAGGGAATGTGCGGTACAGAGAAGGGGAGGCCACGTGCTCTGCAITGAGC 245

QY 252 AATGCAACCTCACAGAGGCGCTGTGTGGCAGTAATGGCAAGACCTTACCTCAACCACT 311
Db 246 AATGCAACCTCACAGAGGCGCTGTGTGGCAGTAATGGCAAGACCTTACCTCAACCACT 305

QY 312 GTGAATGATCGATCGATGCGCTGCTCACTGGATCCAAATCCAGGTCAGTTGATTACGATGAC 371
Db 306 GTGAATTCATAGAGATGCTGCTGCTCACTGGATCCAAATCCAGGTCAGTTGATTGATGAGC 365

QY 372 ACTGCAAGAGAGAAATCCGTAAGTCCATCTGCCAGCCCAAGTCTTTTGTCTATCAGTCCA 431
Db 366 ACTGCAAGAGAGAAAGTCTGCGAGTCCATCTGCCAGCCCAAGTCTTTGCTATCAAGCTA 425

QY 432 ACCTGATGAGTCCGACGTGCATCTCCAGTGCATCTGAGTGCCTGGAAGCTGAGATCATCTCAGATG 491
Db 426 ACCGATGAGTCCGACGCGCGCTCATCTCCAGTGCCTGGAAGCTGAGATCATCTCAGATG 485

QY 492 GCTGCTTCTTAAGGACGACACTACAGTGAATCCTAGACAAATCTTAAAGAACTTTG 551
Db 486 GCTGCTTCTTAAGGACGACACTACAGTGAATCCTAGACAAATCTTAAAGAACTTTG 545

QY 552 ATAAATGCTGATCTCGCTGGAATCCAGTGAATCTGAAAGTTTGTGGAACGAAATGAAA 611
Db 546 ATAAATGCTGATCTCGCTGGAATCCAGTGAATCTGAAAGTTTGTGGAACGAAATGAAA 605

QY 612 CTGCCATCAATATTACAGTATCCAGACCCAGGAGAACAAAGTTGCTTAGGGGACTCT 671
Db 606 CAGCCATCAAGTATCCAGTATCCAGTATCCAGTATCCAGTATCCAGTATCCAGTATCCAGT 665

QY 672 GTGTTGATGCTCTATGTAATGCTGTGTAATAAATGCTGTAATGTAATAAATGCTGTTCCAG 731
Db 666 GTGTTGATGCTCTATGTAATGCTGTGTAATAAATGCTGTAATGTAATAAATGCTGTTCCAG 725

QY 732 AGTTTCTCAAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGT 791
Db 726 AGTTTCTCAAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGT 785

QY 792 ATGAAACGATATCAGATGAGGCTGAGACCGAGGTGAGATGTAAACCGCTGTGTGTGTGCT 851
Db 786 ACCAAACCTATCAGATGAGGCTGAGACCGAGGTGAGATGTAAACCGCTGTGTGTGTGCT 845

QY 852 GTGGAATTTGGTCTGTACAGCCATCACTGTGAGCGGAAGAAATCAGAAAGGGGGCCACAGA 911
Db 846 GTGGCCACTGGGTCTGTACAGCCATCACTGTGAGCGGAAGAAATCAGAAAGGGGGCCACAGA 905

QY 912 CCCAGACAGAGGAGGAGATGACACAGATATGTCCAGAGGCTCCAAAAGATCAGGAAAACAG 971

906 CCCACAGAGGAGGAGACAGAGGATATCTCAGGAATCTCCAGAAAGCCAGGGGACAG 965
972 CTGGAAGACCAAGAGAGTGAAGACCAAGAGATCTAATGAG-GAGGCACAGACAGTCT 1030
966 CAGAAAGACCAAGAGAGTGAAGACCAAGAGATCTAAGAGAGGACAGACCGGTGT 1025
1031 CTGATCCAGCATCTCTCCATCTTCCAGCTCAGCTGAGTTTCAGTATACACAAAGTGTCTCTAC 1090
1026 CCGGAGCCAGGCGCT--CCTCTTACGCTGAGCCAGTACACAC-AGAGTCTGAGC 1081
1091 AGTCGCCAATCACCAGTATTTGCTTATATAGCAATGAGTTTATTTTATTTTATTTT 1150
1082 AATCACAAATACATAGTATTTGCTGTATGAGGAGCAATCTTATTTTGTG---TTTGT 1137
1151 TGCATTAAGGATATGAAGGTGGCTGTAGGAAGGAGGAGGCGCCAGAGCTTCATTCTTA 1210
1138 TGCAATTAAGGAATGAGGTGGCCAGCTAGCGAGGAGGACCAACCTTCACCTGTA 1197
1211 GGAGTCTTTAAGGAACCTGTAAATGCTGCTCTGGGCTGGAGGCTAGTAAGGAACCTG 1270
1198 GGAATGCTTTAAGGAAC--TAAAGACACCTTGGGACGAGAGGCACTTAAGGAACAG 1255
1271 CATCACCATTAAGAGAAACAGACACCCAAATCTGAAC---CTCTTTGAGTTTACTGCA 1326
1256 CATC-GGTTGGCAGAGAGAGAGGAGGAGCTTTGAATGAAGCTTTCTGGGTCACAGCA 1314
1327 TCTGTACAGAGCTGACAGAGTGTCAACATGCGCAGAGAACTTAGCAGGCTGTCCCG 1386
1315 GCTCGAGGAGATACAGGAAGAGCAT-----AGAGAAACATTTGAATGAGCCTG 1364
1387 GGAGCAGAGTTTGGGAAGCTCCACGAGAGGAGAGCTCTCTGCTTCCAGCTCTTTCCA 1446
1365 CTGAGGAAGTGGGGAGCTTTGTAGGAGAGAACCTGCTGCTTT--GACCTTGTCA 1421
1447 TTGCCGTGACATGACAGACTCCAGCATCCAGCATCTCTTGGTCCCAATAAC-TGCCT 1505
1422 CCACTGTACAGATACAGAGCTGACAGAGTGTCTCTCTCTCTCTCTCTCTCTCTCT 1481
1506 CTAGATACATAGCCATCTGCTAGTTAAACAGTGTCCCTCAGACTTGGAGTGTCTT 1565
1482 CTGAAACACAGCGCCCACTAGTTA--CTGTGTCCTCAGCTTGCATGAGTGTCC 1539
1566 GGGAGGTACACCAATGATGACAGATCTGTATATCTTTGAGCCCTTAGCAGCCTAAC 1625
1540 TGGAGGAGGTGTTTAAATGATGACAGACATATGATCTTCAAGCGCATGAGACCTAAC 1599
1626 CAATTTTAAATACATCTTTT-----ACCAAGGTGCTATTTC 1663
1600 CAATTTTAAATACATTTTCTTTTTTTTTTTTGTGTTAACCAAGGTGCTATTTC 1659
1664 TCTGTAAACACATTTTTTTTGGCAAGTTGACTTTATCTTTCAATTTATTAATTA 1723
1660 TCTGTAAAGACTTTTTT---CCAAGCTGACTTCAATCTCAGTTATTAACGTTATTA 1716
1724 TTGT--TTTTTAATTTATTTTCTTCTGACTAGTATTAAGCTTTGTAATTA-TTTTT 1780
1717 TTGTTGTTTTTAAATTTTCAATTTTGTGACTAGATATTAAGCTTTTGTAAATTTTTC 1776
1781 AGTAGTCCCACTTTCATAGTGGG-----AGGAGTTTGGGGTTCTTCTCTGTGACG 1833
1777 ATTAGTCTTACTATTTTCAAGAGTGAAGGTGAAGGGGTTTGGGCATTTTTCAGGTACAG 1836
1834 GGGCTGAATTAACCCAGA-TGCTCCCACTTCCACATACATAGA-----TGCAGCCATA 1887
1837 GGAATCTGTAAACCAACAGCCCATACCTCTGTACATATTAGACCGGTGTGAGTTCGA 1896
1888 GTTGGCC-----CCCTAGTCTTCAGAGTCCA---CTATCTCCAGAGAGCAAGGGTG 1939
1897 GCTGCACCCAACCCAGAGCTTCTAGAAATCAGTCCATGCCAGAGACAGAGGCC 1956
1940 CTTTAGACCTTAAGCCAGGGAAGAGCATCTTCAATAAAACCTTCAAGATCCAAACA-T 1998

Db 1957 CCTCAGACAGAAAGCCACAGGACAAAGCATCTTCTAGACAGCTTGTGATCCAAACAGT 2016
QY 1999 TAATTTGTTTTTATTTATTTCTGAGAAGTT 2027
Db 2017 TAATTTGTTTTTCTTCTGTAAGAGTT 2045
RESULT 14
MUSTSC36A
LOCUS
DEFINITION
MUS musculus TGF-beta-inducible protein (TSC-36) mRNA, complete cds.
ACCESSION
M91380.1 GI:349005
VERSION
TGF-beta-inducible protein.
KEYWORDS
MUS musculus (house mouse)
ORGANISM
MUS musculus
REFERENCE
1 (bases 1 to 2823)
Shibanuma,M., Mashimo,J., Mita,A., Kuroki,T. and Nose,K.
Cloning from a mouse osteoblastic cell line of a set of
transforming-growth-factor-beta 1-regulated genes, one of which
seems to encode a follistatin-related polypeptide
Eur. J. Biochem. 217 (1), 13-19 (1993)
JOURNAL
MEDLINE
PUBMED
94039028
7901004
REFERENCE
2 (bases 1 to 2823)
Nose,K.
AUTHORS
TITLE
JOURNAL
COMMENT
ORIGINAL SOURCE text: Mus musculus cDNA to mRNA.
Location/Qualifiers
1. 2823
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/cell_lines="MC3T3-E1"
/cell_type="osteoblast"
1. 2823
/gene="TSC-36"
80..1000
/gene="TSC-36"
/codon_start=1
/product="TGF-beta-inducible protein"
/protein_id="AAC37633.1"
/db_xref="GI:349006"
/translation="MWKRLALSLVTTLVHGEEPRSKIKANVFCGAGRECAVTE
KGETCLISQCKPKKPVCGSNGKTYLNHCELHRLDACLTKSKIQVDYDGHCKEKKA
SPSASPVCQANDELRRLLIOWLEAIIIPDGFSGSNYSLELDKXFSFNGDSH
LDSSEFLKFFQNETAINITTYADQENKLLRLSLVDALLIELSDENADWKLSPFELK
CLNPSFNPEKKCALEVETVADGAETEDCNRCVCGHWVCTAMTCDGNKQGVQTH
TEBEKTYVOBLOKHQGTAEKTKVNTKEI"
ORIGIN
Query Match 27.8%; Score 1025.4; DB 10; Length 2823;
Best Local Similarity 75.9%; Pred. No. 1.5e-241; Mismatches 406; Indels 93; Gaps 21;
Matches 1572; Conservative 0;
QY 12 CTCCACACTCCGCTTACAGCTTCGCTGCGCGCTCTCTCCCGCCCTCTCCCGCCCGCCCGAGAGACTGG 71
Db 9 CTCCACACTTCGCTTAACTCGCTGCGCCACCTGCGCAGTGTCTCCGAGTCCCGG 68
QY 72 ACCAGACCAAGTGGAAAGCTTCGGCTCGCGCTCGCGCTCGCGCTCGCGCTCGCGCTCGCGCT 131
Db 59 ACCGAGCAGATGTGGAACGATGG-----CTGGCGCTCTCGCTGTGACCATCGCCC 122
QY 132 GGGTCCGCGCGAGAGAGAGTAAAGGAGCAATCAAGATCTGTGCCAATGTGTTTGTG 191
Db 123 TGGTCCACGCGAGGAGGAACTAGAGCAATCAAGATCTGCGCAATGTGTTTGTG 182
QY 192 GAGCGCGCGGGAATGTCAGTTCAGAGAAAGGGAAACCCACTGTCTCTGCAATTGAGC 251

Db	183	GAGCTGCAGGGAAATGTGCCGTACAGAGAAAGGGGAGCCACGTCCTCTGCATTGAGC	242
2Y	252	AATGCAAACTCAAGAGGCGCTGTGTGGCAGTAATGGCAAGACTACTCTCAACCCT	311
2b	243	AATGCAAACTCAAGAGGCGCTGTGTGGCAGTAATGGCAAGACTACTCTCAACCCT	302
2Y	312	GTGAATGCAATCAGAGTGCCTGCTCACTCGATCCAAATCCAGGTTGATTACGATGGAC	371
2b	303	GTGAATTCATAGAGATGCCCTCACTCGATCCAGATCCAGGTTGATTATGATGGC	362
2Y	372	ACTGCAAGAGAAATCCGTAACTTCCATCTCTGCCAGCCCGATGTTGCTATCAGTCCA	431
2b	363	ACTGCAAGAGAAAGAGCTCTGCGAGTCCATCTGCCAGCCCGATGTTGCTATCAGTCCA	422
2Y	432	ACCGTGATGAGTCCCGACGTCGGATCATCCAGTGGCTGGAAGCTGAGATCATCTCAGATG	491
2b	423	ACCGGATGAGCTCCGACGGGCGCTCATCCAGTGGCTGGAAGCTGAGATCATTCAGATG	482
2Y	492	GCTGGTTCTCTAAAGGAGCAACTACAGTGAATCTTAGACAAGTATTTAAGAACTTTG	551
2b	483	GCTGGTTCTCTAAAGGAGCAATACAGTGAAGTCTTAGACAAGTATTTAAGAGCTTTG	542
2Y	552	ATAATGGTGATTTCTCGCTGGACTCCAGTCAAATCTTGAAGTTTGTGGAAACAAGTAA	611
2b	543	ATAATGCGACTCTCACCTGGAATCCAGTGAATCTTGAAGTTTGTGGAGCAAGTAA	602
2Y	612	CTGGCATCAATATTACAACTATCCAGACAGAGAAACAAGATTTGCTTAGGGACTCT	671
Db	603	CAGCCATCAACATACCACTATTGCAATCAGAGAAACAACAACCTGCTCAGAAGCCTCT	662
2Y	672	GTGTGATGCTCTCATTTGAACCTGTGATGAAATGCTGATGGAAACTCAGCTTCCAAG	731
2b	663	GTGTGACGCCCTCATTTGAACCTGTGATGAGAACGCTGACTGGAACCTCAGCTTCCAAG	722
2Y	732	AGTTTCTAAGTGCTCAACCCATCTTCAACCTCTGTGAGAGAGTGTGCCCTGGAGG	791
2b	723	AGTTTCTAAGTGCTCAACCCATCTTCAACCTCTGTGAGAGAGTGTGCCCTGGAGG	782
2Y	792	ATGAAACGTATGCAGATGGAGCTGAGACCGAGGTGGACTCTAAACGCTGTGTGTGTCCT	851
Db	783	TCGAAACCTATGCAGATGGAGCTGAGACTGAGGTGGACTGCAATCGCTGTGTGTTCTCT	842
2Y	852	GTGGAATTTGGTCTGTACAGCATGACCTGTGACGGAAGAAATCAGAAGGGGGCCAGA	911
Db	843	GTGGCCACTGGTCTGCACAGCAATGACCTGTGATGGAAGAAATCAGAAGGGGGTCCAGA	902
2Y	912	CCCAGACAGAGGAGGAGATCACCAGATATGTCAGAGAGCTCCAAAAGCATCAGGAACAG	971
2b	903	CCCACACAGAGGAGGAGACAGACAGATATGTCAGAGAACTCCAGAAGCACCGGCACAG	962
2Y	972	CTGAAAGACCAAGAGAGTGAGCACCAAGAGATCTAATGAG-GAGGCACAGACCAAGTGT	1030
Db	963	CAGAAAAGACCAAGAGGTGAACACCAAGAGATCTAAGAGAGGACAGAGCACCCTGT	1022
QY	1031	CTGGATCCCAGCATCTTCTCCACTTCAGCGCTGAGTTTCAGTATACAAAGTGTCTCTAC	1090
Db	1023	CCGGAGCCCAGCGCT--CCTCTTCAGCGCTGAGCCAGTACACAC-AGAGTCTGCAGC	1078
QY	1091	AGTCGCCAATCACCAGTATTGCTTATATAGCAATGAGTTTTATTGTTTATTGTTTT	1150
Db	1079	AATCACCAATCATAGTATTGCTTGATGGACGCAATCTTATTGTTT--TTTGTTTT	1134
QY	1151	TGCAATAAAGGATATGAAGTGGCTGCTAGGAAGGGAAGGCCACAGCCTTCATTCTTA	1210
Db	1135	TGCAATAAAGGAAATGAGGTTGGCCAGCTAGCGAGGGAAGGCCACAACCTTCACCTGTA	1194
QY	1211	GGAGTGCTTTAAGAGAACTGTAATCGTGTCTGTGGGCTGGAGGCTAGTAAGGAACCTG	1270
Db	1195	GGAATGCTTTAAGGAAC--TAAAGACACCTTGGACGAGAGGCACTAAGGAACAG	1252
QY	1271	CATCACGATTTGAAGAGAAACAGACCAAATCTGAAC----CCTTTTGTAGTTTACTGCA	1326
Db	1253	CATC-GGGTTGGCAGAGGAGCAGAGGAGGTGTAATGAAGCCCTTCTGGGGTCAGACA	1311

RESULT 15
HSM800525 1010 bp mRNA linear PRI 18-FEB-2000
LOCUS Homo sapiens mRNA; cDNA DKFp586O2223 (from clone DKFp586O2223).
DEFINITION AL050228
ACCESSION AL050228.1 GI:4884472
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1010)
AUTHORS Ansong,W., Winkner,U., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
TITLE Direct Submission
JOURNAL Submitted (15-MAY-1999) MIPS, Am Klopferspitz 18a, D-82152
Martinsried, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@kiz-heidelberg.de;
sequenced by EMBL (European Molecular Biology Laboratories,

Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project.
This clone (DKFZp5860223) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at <http://www.mips.biochem.mpg.de/proj/cDNA/>.

FEATURES

source 1..1010
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:DKFZp5860223"
/db_xref="taxon:9606"
/clone="DKFZp5860223"
/issue_type="uterus"
/clone_lib="586 (synonym: hutel). Vector pSport1; host DH10B; sites NorI + SalI/WluI"
/dev_stage="adult"
polyA_signal 991..996
polyA_site 999

ORIGIN

Query Match 27.3%; Score 1008.4; DB 9; Length 1010;
Best Local Similarity 99.9%; Pred. No. 1.9e-237;
Matches 1009; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
2693 TTGACATAAGATCTCATTTTCAGAAAGAGAGATTAAGACCATCAGAGAAATTTATG 2742
1 TTGACATAAGATCTCATTTTCAGAAAGAGAGATTAAGACCATCAGAGAAATTTATG 60
2743 GTTGTAAATGACAGGCAACTGTGAGAACTGTGTGCCAAATAGAAATTCCTCTAGTT 2802
61 GTTGTAAATGACAGGCAACTGTGAGAACTGTGTGCCAAATAGAAATTCCTCTAGTT 120
2803 TTTCTTGTCTCATTTTGAAGAGAGAAATTCCTCTTGTGTAGCAATTCAGCTTTATG 2862
121 TTTCTTGTCTCATTTTGAAGAGAGAAATTCCTCTTGTGTAGCAATTCAGCTTTATG 180
2863 TATCCATCCCATCTAAATCTTCAACTTCCACTTGTTCAGTCTGAAATGAGCTCCCT 2922
181 TATCCATCCCATCTAAATCTTCAACTTCCACTTGTTCAGTCTGAAATGAGCTCCCT 240
2923 GTCCAAAGTGCTTGGAGAACTCAGCAGCAGCCCTTAATCAAGAGTTTACGAGCCCT 2982
241 GTCCAAAGTGCTTGGAGAACTCAGCAGCAGCCCTTAATCAAGAGTTTACGAGCCCT 300
2983 GGACATATGGAGAGGAGGCAAGAGTACACCAATTTGTTAAAGCAAGAACCAAGTGT 3042
301 GGACATATGGAGAGGAGGCAAGAGTACACCAATTTGTTAAAGCAAGAACCAAGTGT 360
3043 CTCTTCACTAGTCAATTAGAACTGGTTATCACTCCAGACTACTCTACCCCTGCAACATTG 3102
361 CTCTTCACTAGTCAATTAGAACTGGTTATCACTCCAGACTACTCTACCCCTGCAACATTG 420
3103 AACTCCCAAGAGCAATCCCATTCCTCTTGTGAGTTCTGAGCTTCTGTGTAATAGGGCA 3162
421 AACTCCCAAGAGCAATCCCATTCCTCTTGTGAGTTCTGAGCTTCTGTGTAATAGGGCA 480
3163 GCTGTCTCTATGCGGTAGAAATCACATGATCTCAGGACCAATTCATGGAAGCTGCTAAATA 3222
481 GCTGTCTCTATGCGGTAGAAATCACATGATCTCAGGACCAATTCATGGAAGCTGCTAAATA 540
3223 GCCTAGTCTGGGAGTCTTCCATAAAAGTTTGGCATGGAGCAAAACAAACAGGATTAACCTA 3282
541 GCCTAGTCTGGGAGTCTTCCATAAAAGTTTGGCATGGAGCAAAACAAACAGGATTAACCTA 600
3283 GGTGTGGTCTTTCAGCCCTCTAAAGCATAGGCTTGTAGCTCAGGCTTCTTGGGCTT 3342
601 GGTGTGGTCTTTCAGCCCTCTAAAGCATAGGCTTGTAGCTCAGGCTTCTTGGGCTT 660
3343 TCTCTGTGTGTAGTTTGTAAACACTATAGCATCTGTGTAATCCAGTGTCCATGGAA 3402
661 TCTCTGTGTGTAGTTTGTAAACACTATAGCATCTGTGTAATCCAGTGTCCATGGAA 720

Search completed: March 11, 2004, 03:35:56
Job time : 9420.52 secs

QY 3403 ACCTTCCACATGCGGTGACTCTGGACTATATCAGTTTGGAAAGCAGGGTTCCTCTGC 3462
Db 721 ACATTCACATGCGGTGACTCTGGACTATATCAGTTTGGAAAGCAGGGTTCCTCTGC 780
QY 3463 CTGCTAAACAAGCCCAACGTCGACAGTCTGAATGTCTTTTCCCTTTACACCTATGTTTTAAG 3522
Db 781 CTGCTAAACAAGCCCAACGTCGACAGTCTGAATGTCTTTTCCCTTTACACCTATGTTTTAAG 840
QY 3523 TAGTCAAACTTCAAGAAACAATCTAAACAAGTTTCTGTTCATATGTTTGTGAACCTG 3582
Db 841 TAGTCAAACTTCAAGAAACAATCTAAACAAGTTTCTGTTCATATGTTTGTGAACCTG 900
QY 3583 TATTTGTATTTAGTAGGCTTCTATATTTGCAATTTAACTTTGTTTGTAACTCCTGATTCCT 3642
Db 901 TATTTGTATTTAGTAGGCTTCTATATTTGCAATTTAACTTTGTTTGTAACTCCTGATTCCT 960
QY 3643 CCTTTTCGGATCTATTGATGAATAAGAAATTAAGTGAATAAGAAATTAAGTGAATAAGAA 3692
Db 961 CCTTTTCGGATCTATTGATGAATAAGAAATTAAGTGAATAAGAAATTAAGTGAATAAGAA 1010

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

M nucleic - nucleic search, using sw model

un on: March 10, 2004, 08:01:57 ; Search time 898.804 Seconds
(without alignments)
17454.974 Million cell updates/sec

itle: US-10-084-817-20

effect score: 3693
sequence: 1 gcatcgccgactccaccct.....ttaaagtgaataaaaaaa 3693

coring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

earched: 3373863 seqs, 2124099041 residues

total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_29Jan04.*

- 1: _geneseqn1980s.*
- 2: _geneseqn1990s.*
- 3: _geneseqn2000s.*
- 4: _geneseqn2001as.*
- 5: _geneseqn2001bs.*
- 6: _geneseqn2002s.*
- 7: _geneseqn2003as.*
- 8: _geneseqn2003bs.*
- 9: _geneseqn2003cs.*
- 10: _geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

result No.	Score	Query Match	Length	DB ID	Description
1	3693	100.0	3693	10	ADE77194 Human CDN
2	1940	52.5	1987	6	ABK64722 Human ben
3	1940	52.5	1987	6	ABN96790 Gene #328
4	1729.8	46.8	2538	7	AB234719 Coding se
5	1729.8	46.8	2538	4	AAL57940 Human pol
6	1678.4	45.4	1755	4	AAL37538 Human mus
7	1678.4	45.4	1755	7	ABX60526 CDNA enco
8	1008.4	27.3	1010	6	ABQ88210 Human bst
9	924.2	25.0	1017	2	AAX39840 Gastric c
10	911.8	24.7	926	2	AAT68831 cDNA for
11	850.2	23.0	1370	9	ABD58526 Toxicity-
12	844	22.9	879	6	AB92989 Human CDN
13	757.2	20.5	881	2	AAX39842 Gastric c
14	729.4	19.8	930	7	AB234644 Coding se
15	729.2	19.7	810	2	AAX39843 Gastric c
16	674.8	18.3	808	2	AAX39844 Gastric c
17	658.4	17.8	879	2	AAX39807 Gastric c
18	657.2	17.8	705	6	ABA92981 Human CDN
19	657	17.8	719	2	AAX39803 Gastric c
20	630.4	17.1	634	4	AAD23560 Human lun
21	630.4	17.1	634	9	ADD66848 Human lun
22	630.4	17.1	634	9	AD688102 Human lun
23	629.4	17.0	708	2	AAX39805 Gastric c

24	619.4	16.8	621	3	AAC65933 Human lun
25	619.4	16.8	621	6	ABL49152 Human lun
26	619.4	16.8	621	6	ABQ92341 Human lun
27	619.4	16.8	621	8	ADA28299 Human lun
28	619.4	16.8	621	9	AD63504 Human lun
29	578.2	15.7	762	2	AAX39804 Gastric c
30	550.2	14.9	730	9	ADD34752 Mouse mit
31	543.2	14.7	589	4	AAD23566 Human lun
32	543.2	14.7	589	9	ADD66854 Human lun
33	543.2	14.7	589	9	AD688108 Human lun
34	520.4	14.1	652	2	AAX39806 Gastric c
35	493.8	13.4	664	2	AAX39802 Gastric c
36	477.2	12.9	788	2	AAX40109 Gastric c
37	440.2	11.9	475	8	ACH37045 Human end
38	439.6	11.9	793	2	AAX39851 Gastric c
39	434.6	11.8	504	8	ACH38219 Human end
40	431.2	11.7	436	8	ACH29732 Human tes
41	427.6	11.6	435	8	ACH21451 Human adu
42	419.6	11.4	477	8	ACH27721 Human adu
43	385.2	10.4	398	3	AAC07466 Human sec
44	356	9.6	356	9	ADD49313 Human lun
45	354.4	9.6	356	9	ADD49659 Human lun

ALIGNMENTS

RESULT 1

AD677194
ID ADE77194 standard; CDNA; 3693 BP.

XX AC ADE77194;

XX DT 29-JAN-2004 (first entry)

XX DE Human cDNA differentially expressed in a liver disorder #271.

XX KW human; ss; gene; liver disorder; hyperlipidaemia; hypertension;

XX KW type II diabetes; tumour; liver; inflammatory disorder;

XX KW immune response disorder; high-throughput screening;

XX KW differential gene expression; gene therapy.

XX XX Homo sapiens.

XX OS US2003108871-A1.

XX FN XX

XX PD 12-JUN-2003.

XX PF 30-JUL-2001; 2001US-00919039.

XX XX 28-JUL-2000; 2000US-0222113P.

XX XX (KASE/) KASER M R.

XX PI Kaser MR;

XX DR WPI; 2004-031227/03.

XX DR P-PSDB; ADE77195.

XX XX Composition comprising several cDNAs that are differentially expressed in

XX XX treated human C3A liver cell cultures, useful for treating liver

XX XX disorders.

XX XX Claim 1; SEQ ID NO 359; 41pp; English.

XX XX The invention relates to a composition comprising several cDNAs that are differentially expressed in a liver disorder. The composition is useful for treating liver disorder such as hyperlipidaemia, hypertension, type II diabetes, tumours of the liver and disorders of the inflammatory and immune response. The composition is useful for a high-throughput method of screening several molecules or compounds to identify a ligand which specifically binds a cDNA. A protein encoded by the cDNA is useful for a high-throughput method for using a protein to screen several molecules or

1921 TGCCAGAGGAGCAAGGGTGCCTTAGACCTAAGCCAGGGAAGCATCTTCATAAAAA 1980
 1981 CTTTCAAGATCCAAACATTAATTTGTTTTATTTATTTCTGAGAGTTGAGGCAATCAGT 2040
 1981 CTTTCAAGATCCAAACATTAATTTGTTTTATTTCTGAGAGTTGAGGCAATCAGT 2040
 2041 ATTCCCAAGGATCGGACCAAGGCGACCAAGCGCTTAGGATATCCAGCGCTACCAAT 2100
 2041 ATTCCCAAGGATCGGACCAAGGCGACCAAGCGCTTAGGATATCCAGCGCTACCAAT 2100
 2101 ATGCTCATTGCACTAATAGAGGGGTGAGTTGGCCCTGCTCTCTCTTTCTGAGCCTC 2160
 2101 ATGCTCATTGCACTAATAGAGGGGTGAGTTGGCCCTGCTCTCTCTTTCTGAGCCTC 2160
 2161 AGTTTCTCAGTGCAGTGGTGAAGATGCACAACTTTTGAATGATAAGTTATATAATTC 2220
 2161 AGTTTCTCAGTGCAGTGGTGAAGATGCACAACTTTTGAATGATAAGTTATATAATTC 2220
 2221 TGTGGTTCCTGATCATTTGGTCCAGAGGGGAGATAGGTTCCCTGATTTTCCCTTCTCT 2280
 2221 TGTGGTTCCTGATCATTTGGTCCAGAGGGGAGATAGGTTCCCTGATTTTCCCTTCTCT 2280
 2281 ATAGATAAATAGAAATCTTGTACTAGAAACAAGAAATGTCAGATGGCCCAAAAACAAGATG 2340
 2281 ATAGATAAATAGAAATCTTGTACTAGAAACAAGAAATGTCAGATGGCCCAAAAACAAGATG 2340
 2341 ACCAGATTTGATCTCAGCCTGATGACCCCTACAGTGTGCTATGATAGTGCCTCATG 2400
 2341 ACCAGATTTGATCTCAGCCTGATGACCCCTACAGTGTGCTATGATAGTGCCTCATG 2400
 2401 GGTAAAGCAGAGAGAGTGGGAAAGAGAACCCACCCACTCTCTTCATTTTGCATTT 2460
 2401 GGTAAAGCAGAGAGAGTGGGAAAGAGAACCCACCCACTCTCTTCATTTTGCATTT 2460
 2461 CATGTTAACTCCGCTGGAATAGAAATAGAAAGCANTCCCTTAGATGAGGATAAAGAAAG 2520
 2461 CATGTTAACTCCGCTGGAATAGAAATAGAAAGCANTCCCTTAGATGAGGATAAAGAAAG 2520
 2521 TTTCAGATTCACAGGGGGAGAAATGGAGATTTAATCCTAAATCTGACCTGGGGAG 2580
 2521 TTTCAGATTCACAGGGGGAGAAATGGAGATTTAATCCTAAATCTGACCTGGGGAG 2580
 2581 GTGAGTCATTTACAGTTAGTCTGTGCTTTTCGACTTCTGTGATTAACCCCACTCAC 2640
 2581 GTGAGTCATTTACAGTTAGTCTGTGCTTTTCGACTTCTGTGATTAACCCCACTCAC 2640
 2641 TACCTGTTCAGATGCAATTTGGAATACCAAGATTAATCCTTGACATAGATCTCAT 2700
 2641 TACCTGTTCAGATGCAATTTGGAATACCAAGATTAATCCTTGACATAGATCTCAT 2700
 2701 TGCAGAAAGCAGATTAAGACCATCAGAGGAAATTTATTTAGTGTGTAATGCACAGGCAA 2760
 2701 TGCAGAAAGCAGATTAAGACCATCAGAGGAAATTTATTTAGTGTGTAATGCACAGGCAA 2760
 2761 CTGTGAGAACTGTGTGCCAAATAGAAATCCTCTGATTTTCTGTGTTCTCATTTGA 2820
 2761 CTGTGAGAACTGTGTGCCAAATAGAAATCCTCTGATTTTCTGTGTTCTCATTTGA 2820
 2821 AAGGAGAAATTCACATTTGTTAGCAATTCAGCTTTATGATATCCATCCCATCAAAA 2880
 2821 AAGGAGAAATTCACATTTGTTAGCAATTCAGCTTTATGATATCCATCCCATCAAAA 2880
 2881 ACTCTTCAAACTCCACTGTGCTGAGATGAGATGAGCTCCCTGTCAGAGTGCCTTGAGA 2940
 2881 ACTCTTCAAACTCCACTGTGCTGAGATGAGATGAGCTCCCTGTCAGAGTGCCTTGAGA 2940
 2941 ACTCAGCAGCAGCCTTAACTAAAGGTTTACAGCCCTTGACACTATGGAGGAGG 3000
 2941 ACTCAGCAGCAGCCTTAACTAAAGGTTTACAGCCCTTGACACTATGGAGGAGG 3000
 3001 GCAAGAGTACCAATTTGTTTAAAGAGAGAAACACAGTGTCTCTTCACTAGTCATTTA 3060

Db 3001 GCAAGAGTACCAATTTGTTTAAAGAGAGAAACACAGAGTGTCTCTTCACTAGTCATTTA 3060
 Qy 3061 GAACATGTTTATCATCCAGAGCTACTTACCTGCAACATTTGAATCCCAAGAGCAAAATC 3120
 Db 3061 GAACATGTTTATCATCCAGAGCTACTTACCTGCAACATTTGAATCCCAAGAGCAAAATC 3120
 Qy 3121 CACATTTCTCTGAGTTCTGCAGCTTCTGTAAATAGGCGAGCTGTCGTCTATGCCGTA 3180
 Db 3121 CACATTTCTCTGAGTTCTGCAGCTTCTGTAAATAGGCGAGCTGTCGTCTATGCCGTA 3180
 Qy 3181 GAATCACATGATCTGAGGACCATTCATGGAAGCTGTAAATAGCCTAGTCTGGGAGTCT 3240
 Db 3181 GAATCACATGATCTGAGGACCATTCATGGAAGCTGTAAATAGCCTAGTCTGGGAGTCT 3240
 Qy 3241 TCCATAAAGTTTTGCATGGAGCAAAACAGGANTTAACTAGGTTTGGTTCCITCAGCC 3300
 Db 3241 TCCATAAAGTTTTGCATGGAGCAAAACAGGANTTAACTAGGTTTGGTTCCITCAGCC 3300
 Qy 3301 CTCTAAAAGCATAGGGCTTAGCCTGSCAGGCTTCTTGGGCTTCTCTGTGTGTAGTTT 3360
 Db 3301 CTCTAAAAGCATAGGGCTTAGCCTGSCAGGCTTCTTGGGCTTCTCTGTGTGTAGTTT 3360
 Qy 3361 TGTAAACACTATGATCTGTTAAGATCCAGTGTCCATGGAACCTTCCACATGCCGTG 3420
 Db 3361 TGTAAACACTATGATCTGTTAAGATCCAGTGTCCATGGAACCTTCCACATGCCGTG 3420
 Qy 3421 ACTCTGGACTATATCAGTTTTTGGAAAGCAGGGTTCCTCTGCTGCTCTTAAAGAGCC 3480
 Db 3421 ACTCTGGACTATATCAGTTTTTGGAAAGCAGGGTTCCTCTGCTGCTCTTAAAGAGCC 3480
 Qy 3481 GGACAGTCTGAATGCTTCTTCTTACACCTATGTTTAAAGTAGTCAAACTTCAAGAAA 3540
 Db 3481 GGACAGTCTGAATGCTTCTTCTTACACCTATGTTTAAAGTAGTCAAACTTCAAGAAA 3540
 Qy 3541 CAATCTAAACAAAGTTTCTGTTGCATATGTTTGTGAACCTGTATTTAGTAGGCG 3600
 Db 3541 CAATCTAAACAAAGTTTCTGTTGCATATGTTTGTGAACCTGTATTTAGTAGGCG 3600
 Qy 3601 TTCTATATGCAATTAACCTGTTTGTAACTCTGATCTCTTCTTTCGGATACTATTG 3660
 Db 3601 TTCTATATGCAATTAACCTGTTTGTAACTCTGATCTCTTCTTTCGGATACTATTG 3660
 Qy 3661 ATGAATAAAGAAATTAAGTGAAAAA 3693
 Db 3661 ATGAATAAAGAAATTAAGTGAAAAA 3693

RESULT 2
 ABK64722
 ID ABK64722 standard; DNA; 1987 BP.
 XX
 AC ABK64722;
 XX
 DT 18-JUN-2002 (first entry)
 XX
 DE Human benign prostatic hyperplasia gene #617.
 XX
 KW Human; benign prostatic hyperplasia; BPH; prostate cancer; gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200212440-A2.
 XX
 PD 14-FEB-2002.
 XX
 PP 07-AUG-2001; 2001WO-US024708.
 XX
 PR 07-AUG-2000; 2000US-0223323P.
 XX
 PR 05-JUN-2001; 2001US-00873319.
 XX
 PA (GENE-) GENE LOGIC INC.
 PA (NIB) JAPAN TOBACCO INC.
 XX

PI Munger WE, Kulkarni P, Getzenberg RH, Waga I, Yamamoto J;
CX WFI; 2002-257476/30.
CX Identifying drugs for and diagnosing benign prostatic hyperplasia, by
CX detecting expression levels of one or more genes in prostate cells from
CX patient that are differentially regulated compared to normal prostate
CX cells.
CX
CX Disclosure; Page 328-329; 444pp; English.
CX
CX The invention relates to a method of diagnosing (I) the onset or
CX progression of benign prostatic hyperplasia (BPH), or screening (II) for
CX or identifying an agent that modulates the onset or progression of BPH.
CX The method is based on changes in gene expression in BPH tissue isolated
CX from patients exhibiting different clinical states of prostate
CX hyperplasia as compared to normal prostate tissue. (I) comprises
CX detecting the expression levels of one or more genes in prostate cells
CX from the subject that are differentially regulated compared to normal
CX prostate cells. (II) comprises preparing a first gene expression profile
CX of BPH cells or BPH-like cell population, exposing the cells to the
CX agent, preparing a second gene expression profile of the agent exposed
CX cells, and comparing the first and second gene expression profiles. (I)
CX is useful for diagnosing the onset or progression of BPH. (II) is useful
CX for identifying an agent that modulates the onset or progression of BPH.
CX The methods are useful to present information identifying the expression
CX level in a tissue or cells, by comparing the expression level of genes
CX given in the specification in the tissue or cells to the level of
CX expression of gene in the database, and displaying the expression levels
CX of at least one gene in the tissue or cell sample compared to the
CX expression level in BPH. Agents using (II) are useful for treating BPH or
CX prostate cancer. ABK64106-ABK64860 represent human benign prostatic
CX hyperplasia gene sequences of the invention
CX
CX Sequence 1987 BP; 522 A; 482 C; 498 G; 485 T; 0 U; 0 Other;
CX
CX Query Match 52.5%; Score 1940; DB 6; Length 1987;
CX Best Local Similarity 99.7%; Pred. No. 0;
CX Matches 1954; Conservative 0; Mismatches 5; Indels 1; Gaps 1;
CX
CX 12 CTCGCCCTCCGTTACAGTCGTCGCGCGCTGTCGCGCGCGCGCGCGCGCGCGCGCGCTG 71
CX 21 CTCGCCCTCCGTTACAGTCGTCGCGCGCTGTCGCGCGCGCGCGCGCGCGCGCGCGCTG 80
CX 72 ACCAGACCGAGTGGAAAGCTGCTCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTG 131
CX 81 ACCAGACCGAGTGGAAAGCTGCTCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTG 140
CX 132 GGGTCCGCGCGGAGAGCTAAGGAGCAAAATCCAAGATCTGTGCCAATGTGTTTGTG 191
CX 141 GGGTCCGCGCGGAGAGCTAAGGAGCAAAATCCAAGATCTGTGCCAATGTGTTTGTG 200
CX 192 GAGCGCGCGGAGATGTCAGTCACAGAGAGGGAAGCCACCTGCTCTCATTTGAGC 251
CX 201 GAGCGCGCGGAGATGTCAGTCACAGAGAGGGAAGCCACCTGCTCTCATTTGAGC 260
CX 252 AATGCAAACTCACAAGAGCGCTGTGTGCGCAGTAATGCAAGACCTACCTCAACCACT 311
CX 261 AATGCAAACTCACAAGAGCGCTGTGTGCGCAGTAATGCAAGACCTACCTCAACCACT 320
CX 312 GTGACATGCAATGAGATGCTGCTCACTGATGCAAAATCCAAGTGTGATGAGTGGAC 371
CX 321 GTGACATGCAATGAGATGCTGCTCACTGATGCAAAATCCAAGTGTGATGAGTGGAC 380
CX 372 ACTGCAAGAGAGAAATCCGTAAGTCCATCTGCGCAGCCAGTGTGTTGCTATCAGTCCA 431
CX 381 ACTGCAAGAGAGAAATCCGTAAGTCCATCTGCGCAGCCAGTGTGTTGCTATCAGTCCA 440
CX 432 ACCGTGATGAGTCCGAGTGGATCATCCAGTGGCTGGAGCTGAGATCATCTCAGATG 491
CX 441 ACCGTGATGAGTCCGAGTGGATCATCCAGTGGCTGGAGCTGAGATCATCTCAGATG 500
CX 492 GCTGGTCTCTAAGGCGAGCACTACAGTGAATCTCTAGACAAAGTATTTAAGAACTTTG 551

DB 501 GCTGGTCTCTAAGGCGAGCACTACAGTGAATCTAGACAAAGTATTTAAGAACTTTG 560
QY 552 ATAATGAGTATCTCGCTCGAATCTCAGTGAATCTCAGTGAATCTCAGTGAATCTCAGTGA 611
DB 561 ATAATGAGTATCTCGCTCGAATCTCAGTGAATCTCAGTGAATCTCAGTGAATCTCAGTGA 620
QY 612 CTGCCATCAATATTACAACTGATCCAGACGAGGAGAACTCAAGTGTGTTAGGGAGCTCT 671
DB 621 CTGCCATCAATATTACAACTGATCCAGACGAGGAGAACTCAAGTGTGTTAGGGAGCTCT 680
QY 672 GTGTTGATGCTCTCATTTGAATCTGATGAAATCTGATGAAATCTGATGAAATCTGATGAA 731
DB 681 GTGTTGATGCTCTCATTTGAATCTGATGAAATCTGATGAAATCTGATGAAATCTGATGAA 740
QY 732 AGTTTCTCAAGTGTCTCAACCCATCTTCAACCCCTCTGAGAAAGTGTGCTGAGG 791
DB 741 AGTTTCTCAAGTGTCTCAACCCATCTTCAACCCCTCTGAGAAAGTGTGCTGAGG 800
QY 792 ATGAAACGATATGAGATGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 851
DB 801 ATGAAACGATATGAGATGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 860
QY 852 GTGAAATTCGGTCTGTACAGCCATGACCTGTGAGAAAGTGAAGTGAAGTGAAGTGAAGTGA 911
DB 861 GTGAAATTCGGTCTGTACAGCCATGACCTGTGAGAAAGTGAAGTGAAGTGAAGTGAAGTGA 920
QY 912 CCCAGACAGAGGAGAGATGACAGATATGTCCAGGAGTCTCCAAAGCATCAGGAAACAG 971
DB 921 CCCAGACAGAGGAGAGATGACAGATATGTCCAGGAGTCTCCAAAGCATCAGGAAACAG 980
QY 972 CTGAAAGACCCAGAGATGAGACCAAGAGATCTAATCAGGAGGACAGACAGTGTGTC 1031
DB 981 CTGAAAGACCCAGAGATGAGACCAAGAGATCTAATCAGGAGGACAGACAGTGTGTC 1040
QY 1032 TGGATCCAGCATCTTCTCCACTTCCAGCTGAGTTCAGTATACAAAGTGTGTCATCA 1091
DB 1041 TGGATCCAGCATCTTCTCCACTTCCAGCTGAGTTCAGTATACAAAGTGTGTCATCA 1100
QY 1092 GTCCCAAAATCACCAGTATTTGCTTATATAGCAATGAGTATTTATTTTATTTATTTT 1151
DB 1101 GTCCCAAAATCACCAGTATTTGCTTATATAGCAATGAGTATTTATTTTATTTTATTTT 1160
QY 1152 GCAATAAAGATATGAAGTGTGGTGTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1211
DB 1161 GCAATAAAGATATGAAGTGTGGTGTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1220
QY 1212 GAGTGTCTTAAGAGAAATCTGAAATGTGCTCTGCGGCTGAGGCTAGTAAAGAACTGC 1271
DB 1221 GAGTGTCTTAAGAGAAATCTGAAATGTGCTCTGCGGCTGAGGCTAGTAAAGAACTGC 1280
QY 1272 ATCAGATTTGAAAGAGAAACAGACCCAAATCTGAACTCTTTTGAAGTGTACTGCTATCTGT 1331
DB 1281 ATCAGATTTGAAAGAGAAACAGACCCAAATCTGAACTCTTTTGAAGTGTACTGCTATCTGT 1340
QY 1332 CAGCAGGCTGAGGAGTGCACAGATGCCAGAGAGAACTTAGCAGGCTGTCCTCCGAGG 1391
DB 1341 CAGCAGGCTGAGGAGTGCACAGATGCCAGAGAGAACTTAGCAGGCTGTCCTCCGAGG 1400
QY 1392 AGAGTTTGGGAAGCTCCAGGAGAGAACTCTGCTTCCAGGCTCTTTCCATTTGCC 1451
DB 1401 AGAGTTTGGGAAGCTCCAGGAGAGAACTCTGCTTCCAGGCTCTTTCCATTTGCC 1460
QY 1452 GTGAGATGACAGACCTCCAGCATCCAGATCTCTTTGGTCCCAATTAATCTGCTCTAGAT 1511
DB 1461 GTGAGATGACAGACCTCCAGCATCCAGCATCTCTTTGGTCCCAATTAATCTGCTCTAGAT 1520
QY 1512 ACATAGCCATCTAGTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAA 1571
DB 1521 ACATAGCCATCTAGTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAA 1580
QY 1572 GTACACCCAAATGATGAGATCTTTGATATCTTTGAGCCCTTAGCCACTTAACCAATTT 1631

Db 1581 GTACACCAAAATGATGAGATACCTTGATACCTTTAGAGCCCTTAGGACCTAACCAAAATT 1640
2y 1632 TTAATAAATACCTTTTACAAAGGTGCTATTTCTCTGTAAACACTTTTTTTTGGCAAGTT 1691
Db 1641 TTAATAAATACCTTTTACCAAGGTGCTATTTCTCTGTAAACAC-TTTTTTGGCAAGTT 1699
2y 1692 GACTTATATCTTCAAT 1751
Db 1700 GACTTATATCTTCAAT 1759
2y 1752 TAGTATTAAGCTTTTGTAAATATTTTTCAGTAGTCCACCACTTCATAGGTGGAAGGAG 1811
Db 1760 TAGTATTAAGCTTTTGTAAATATTTTTCAGTAGTCCACCACTTCATAGGTGGAAGGAG 1819
2y 1812 TTTGGGGTTCTTCTGTTGAGGGGTGAAATAACCCAGATGTCCTCCACCCCTGCCACATA 1871
Db 1820 TTTGGGGTTCTTCTGTTGAGGGGTGAAATAACCCAGATGTCCTCCACCCCTGCCACATA 1879
2y 1872 CTAGATGACCCATAGTTGGCCCCCTAGCTTCCAGCAGTCCACTATCTGCCAGAGGAG 1931
Db 1880 CTAGATGACCCATAGTTGGCCCCCTAGCTTCCAGCAGTCCACTATCTGCCAGAGGAG 1939
2y 1932 CAAGGGTGCCTTAGACCTAAGCCAGGGGAAGAAGCATCTT 1971
Db 1940 CAAGGGTGCCTTAGACCTAAGCCAGGGGAAGAAGCATCTT 1979

RESULT 3

ABN96790
ID ABN96790 standard; DNA; 1987 BP.

CC AC ABN96790;

CT 13-AUG-2002 (first entry)

DE Gene #3288 used to diagnose liver cancer.

W Gene; liver cancer; db; hepatocellular carcinoma; hepatotropic;
metastatic liver tumour; cytostatic; expression profile; disease state;
disease progression; drug toxicity; drug efficacy; drug metabolism.

X Homo sapiens.

X WO200229103-A2.

X 11-APR-2002.

X 02-OCT-2001; 2001WO-US030589.

X 02-OCT-2000; 2000US-0237054P.

X (GENE-) GENE LOGIC INC.

X Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;

X WPI; 2002-426119/45.

X Diagnosing and detecting the progression of liver cancer, hepatocellular
carcinoma or metastatic liver tumor in a patient, involves detecting the
level of expression of two or more genes in a liver tissue sample.

X Claim 1; SEQ ID NO 3288; 298pp; English.

X The invention relates to a novel method for diagnosing and detecting the
progression of liver cancer, hepatocellular carcinoma or metastatic liver
tumor in a patient, and differentiating metastatic liver cancer from
hepatocellular carcinoma in a patient, involving detecting the level of
expression of two or more genes represented in ABN93503-ABN97455 in a
tissue sample. The method of the invention has hepatotropic, and
cytostatic activity. The method is useful for diagnosing and detecting
the progression of liver cancer, hepatocellular carcinoma and metastatic
liver carcinoma in a patient. The method is useful for identifying
expression profiles which serve as useful diagnostic markers as well as

CC markers that can be used to monitor disease states, disease progression,
drug toxicity, drug efficacy and drug metabolism. Note: The sequence data
for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 1987 BP; 522 A; 482 C; 498 G; 485 T; 0 U; 0 Other;

Query Match 52.5%; Score 1940; DB 6; Length 1987;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 1954; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

Qy 12 CTCGCCACCTCGCTTACAGCTCGCTGCGCGCGCTCTCGCGCGCGCGCGCGCGCGCGCGCGCGCTGG 71

Db 21 CTCGCCACCTCGCTTACAGCTCGCTGCGCGCGCTCTCGCGCGCGCGCGCGCGCGCGCGCGCGCTGG 80

Qy 72 ACCAGACACCATGTGAAACGCTGCTCGCGCTCGCGCTCGCGCTCGCGCTCGCGCTCGCGCTCGCGCT 131

Db 81 ACCAGACACCATGTGAAACGCTGCTCGCGCTCGCGCTCGCGCTCGCGCTCGCGCTCGCGCTCGCGCT 140

Qy 132 GGGTCCGGCGCGAGAGAGCTAAGAGGCAATCCAGAGATCTGTGCCAATGTGTTTGTG 191

Db 141 GGGTCCGGCGCGAGAGAGCTAAGAGGCAATCCAGAGATCTGTGCCAATGTGTTTGTG 200

Qy 192 GAGCGCGCGGGAAATGTGCAGTCAACAGAGAAAGGGAAACCCACCTGTCTCTGCATTGAGC 251

Db 201 GAGCGCGCGGGAAATGTGCAGTCAACAGAGAAAGGGAAACCCACCTGTCTCTGCATTGAGC 260

Qy 252 AATGCAAACTCTCAAGAGGCTGTGTGCGAGTAATGCGAAGACCTACTCTCAACCACT 311

Db 261 AATGCAAACTCTCAAGAGGCTGTGTGCGAGTAATGCGAAGACCTACTCTCAACCACT 320

Qy 312 GTGAACCTGCATCGAGATGCTGCTCACTGCATCCAAATCCAGGTTGATTACGATGAC 371

Db 321 GTGAACCTGCATCGAGATGCTGCTCACTGCATCCAAATCCAGGTTGATTACGATGAC 380

Qy 372 ACTGCAAGAGAAATCCGTAACTCATCTCCAGCCAGTTGTTGCTATCAGTCCA 431

Db 381 ACTGCAAGAGAAATCCGTAACTCATCTCCAGCCAGTTGTTGCTATCAGTCCA 440

Qy 432 ACCGTGATGAGCTCCGACGCTGCATCATCCAGTGGCTGGAAGCTGAGATCATCTCAGATG 491

Db 441 ACCGTGATGAGCTCCGACGCTGCATCATCCAGTGGCTGGAAGCTGAGATCATCTCAGATG 500

Qy 492 GCTGTTCTCTTAAAGGAGCAACTACAGTGAATCTCTAGACAAAGTATTTTAAAGACTTTG 551

Db 501 GCTGTTCTCTTAAAGGAGCAACTACAGTGAATCTCTAGACAAAGTATTTTAAAGACTTTG 560

Qy 552 ATAATGTTGATTTCTGCGCTGAGCTCCAGTGAATCTCTGAAGTTTGTGGAACAGATGAAA 611

Db 561 ATAATGTTGATTTCTGCGCTGAGCTCCAGTGAATCTCTGAAGTTTGTGGAACAGATGAAA 620

Qy 612 CTGCCATCAATATTACAACTGATCCAGACAGGAGAAACAAAGTTCTTTAGGGGACTCT 671

Db 621 CTGCCATCAATATTACAACTGATCCAGACAGGAGAAACAAAGTTCTTTAGGGGACTCT 680

Qy 672 GTGTTGATGCTCTCAATTGAATCTGTGATGAAATGCTGATGAAACTCAGCTTCCAG 731

Db 681 GTGTTGATGCTCTCAATTGAATCTGTGATGAAATGCTGATGAAACTCAGCTTCCAG 740

Qy 732 AGTTTCTCAAGTGCCTCAACCCATCTTTCAACCTCTGAGAGAGAGTGTGCCCTGGAGG 791

Db 741 AGTTTCTCAAGTGCCTCAACCCATCTTTCAACCTCTGAGAGAGAGTGTGCCCTGGAGG 800

Qy 792 ATGAAAAGTATGAGATGAGCTGAGACCGAGGTGAGCTGTAAACCGCTGTGTCTGTCCT 851

Db 801 ATGAAAAGTATGAGATGAGCTGAGACCGAGGTGAGCTGTAAACCGCTGTGTCTGTCCT 860

Qy 852 GTGAAAATTTGGTCTGTACAGCATGACCTGTGCGGAAAGAAATCAGAGGGGGCCGAGA 911

Db 861 GTGAAAATTTGGTCTGTACAGCATGACCTGTGCGGAAAGAAATCAGAGGGGGCCGAGA 920

Qy 912 CCCAGACAGAGGAGGATGACAGATATGTCAGGAGCTTCCAAAGCATCAGGAAAACAG 971

Db 921 CCCAGACAGAGGAGGAGATGACAGATATGTCCAGGAGCTCCAAAGCATCAGAAACAG 980
 Qy 972 CTGAAAGACCAAGAGAGTGGAGCCAAAGAGATCTAATAGGAGGACACAGACAGTGC 1031
 Db 981 CTGAAAGACCAAGAGAGTGGAGCCAAAGAGATCTAATAGGAGGACACAGACAGTGC 1040
 Qy 1032 TGGATCCAGCATCTTCTCCACTTCAGCGTGTAGTTCAGTATACAAAGTGTCTGTACA 1091
 Db 1041 TGGATCCAGCATCTTCTCCACTTCAGCGTGTAGTTCAGTATACAAAGTGTCTGTACA 1100
 Qy 1092 GTCGCCAAATCACCAGTATTTGCTTATATAGCAAGAGTATTTATTTGTTATTTGTTT 1151
 Db 1101 GTCGCCAAATCACCAGTATTTGCTTATATAGCAAGAGTATTTATTTGTTATTTGTTT 1160
 Qy 1152 GCAATAAGGATATGAAGTGGCTGTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1211
 Db 1161 GCAATAAGGATATGAAGTGGCTGTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1220
 Qy 1212 GAGTGTCTTAAGAGAACTGTGTAATGTGTCTGTGGGCTGGAGGCTAGTAAGAGAACTGC 1271
 Db 1221 GAGTGTCTTAAGAGAACTGTGTAATGTGTCTGTGGGCTGGAGGCTAGTAAGAGAACTGC 1280
 Qy 1272 ATCAGGATGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1331
 Db 1281 ATCAGGATGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1340
 Qy 1332 CAGCAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1391
 Db 1341 CAGCAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1400
 Qy 1392 AGAGGTTGGGAGGCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1451
 Db 1401 AGAGGTTGGGAGGCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1460
 Qy 1452 GTCAGCATGACAGACCTCCAGCATCTCTTGGTCCCAATAAATGCTCTAGAT 1511
 Db 1461 GTCAGCATGACAGACCTCCAGCATCTCTTGGTCCCAATAAATGCTCTAGAT 1520
 Qy 1512 ACATAGCCATCTCTAGTATTAACAGAGTGTCTTGTGAGGAGTGTCTGGAGG 1571
 Db 1521 ACATAGCCATCTCTAGTATTAACAGAGTGTCTTGTGAGGAGTGTCTGGAGG 1580
 Qy 1572 GTACACCAATATGAGATCTGTATATCTTTGAGGCTGTAGGAGGAGGAGGAGGAGGAGGAG 1631
 Db 1581 GTACACCAATATGAGATCTGTATATCTTTGAGGCTGTAGGAGGAGGAGGAGGAGGAGGAG 1640
 Qy 1632 TTAATAATCTTTTACAAAGGAGTATTTCTGTGTAACACTTTTGTGGAAGT 1691
 Db 1641 TTAATAATCTTTTACAAAGGAGTATTTCTGTGTAACACTTTTGTGGAAGT 1699
 Qy 1692 GACTTATCTCAATTAT 1751
 Db 1700 GACTTATCTCAATTAT 1759
 Qy 1752 TAGGTATTAAGCTTTGTAATATATTTTCTAGTATCCACACTTCATAGTGGAGGAG 1811
 Db 1760 TAGGTATTAAGCTTTGTAATATATTTTCTAGTATCCACACTTCATAGTGGAGGAG 1819
 Qy 1812 TTTGGGTTCTTCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1871
 Db 1820 TTTGGGTTCTTCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1879
 Qy 1872 CTAGATGAGCCCATAGTGTGGCCCCCTAGCTTCCAGCAGTCCACTATCTGCCAGGAGGAG 1931
 Db 1880 CTAGATGAGCCCATAGTGTGGCCCCCTAGCTTCCAGCAGTCCACTATCTGCCAGGAGGAG 1939
 Qy 1932 CAAGGTTGCTTACACCTAAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1971
 Db 1940 CAAGGTTGCTTACACCTAAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1979

RESULT 4

ABZ34719
 ID ABZ34719 standard, cDNA; 2538 BP.
 XX
 AC ABZ34719;
 XX
 DT 04-FEB-2003 (first entry)
 XX
 DE Coding sequence SEQ ID 77, upregulated in osteogenesis.
 XX
 KW Osteopathic; osteogenesis modulator; gene therapy; osteogenesis;
 XX osteoporosis; bone disease; downregulator; human; ss.
 XX
 OS Homo sapiens.
 XX
 PN W0200281745-A2.
 XX
 PD 17-OCT-2002.
 XX
 PF 05-APR-2002; 2002WO-IB002211.
 XX
 PR 05-APR-2001; 2001US-0281400P.
 XX
 PA (AVET) AVENTIS PHARMA SA.
 XX
 PI Garcia T, Roman Roman S, Baron R, Call K, Theilhaber J;
 PI Connolly T, Jackson A, Bushnell SE, Rawadi G;
 XX
 DR WPI; 2003-058567/05.
 XX
 PT Novel isolated nucleic acid upregulated/downregulated in osteogenesis,
 XX useful for bone disease therapy in subject.
 PT
 PS Claim 26; Page 115-116; 237pp; English.
 XX
 CC The present invention relates to novel nucleotide sequences, which are
 CC differentially expressed in models of osteogenesis upon being put in
 CC contact with a stimulator of osteogenesis. The present sequence is one
 CC such sequence. This sequence can be used for diagnosing osteoporosis/bone
 CC disease in a patient, promoting osteogenesis and/or preventing
 CC osteoporosis/bone disease. The present sequence encodes a secreted
 CC protein
 XX
 SQ Sequence 2538 BP; 707 A; 584 C; 595 G; 652 T; 0 U; 0 Other;
 Query Match 46.8%; Score 1729.8; DB 7; Length 2538;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1742; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
 QY 12 CTCCACCTCCGCTTACAGCTCGTCCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 71
 Db 6 CTCCACCTCCGCTTACAGCTCGTCCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 65
 QY 72 ACCAGACACAGATGTGGAACGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 131
 Db 66 ACCAGACACAGATGTGGAACGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 125
 QY 132 GGGTCCGCGCCGAGGAGAGCTAAGGAGCAATCAAGATCTGTGCCAATGTGTTTGTG 191
 Db 126 GGGTCCGCGCCGAGGAGAGCTAAGGAGCAATCAAGATCTGTGCCAATGTGTTTGTG 185
 QY 192 GAGCGCGCGGGAATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 251
 Db 186 GAGCGCGCGGGAATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 245
 QY 252 AATGCAAACTCTCAAGAGGCTGTGTGTGGCAGTAATGGAAGACCTTCAACCACT 311
 Db 246 AATGCAAACTCTCAAGAGGCTGTGTGTGGCAGTAATGGAAGACCTTCAACCACT 305
 QY 312 GTGAATGCTGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 371
 Db 306 GTGAATGCTGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 365
 QY 372 ACTGCAAAAGAGAAATCCGTAAGTCCATCTGCCAGCCCACTGTTGTTGCTATCAGTCCA 431

366 ACTGCAAGAGAGAAATCGGTAACTGCTCCAGCCAGTGTGTGCTATCAGTCCA 425
 432 ACCGTGATGAGTCCGACGTCGATCATCATCAGTGGCTGGAGCTGAGATCATCTCAGATG 491
 426 ACCGTGATGAGTCCGACGTCGATCATCATCAGTGGCTGGAGCTGAGATCATCTCAGATG 485
 492 GCTGGTTCTCTAAGGAGGAGCAACTACAGTGAATCTTAGCAAGATTTTAAAGAACTTTG 551
 486 GCTGGTTCTCTAAGGAGGAGCAACTACAGTGAATCTTAGCAAGATTTTAAAGAACTTTG 545
 552 ATATGTTGATCTCTGCTGCTGAGTCCAGTCAATCTCTGAAATTTTGGACACAGATGAAA 611
 546 ATATGTTGATCTCTGCTGCTGAGTCCAGTCAATCTCTGAAATTTTGGACACAGATGAAA 605
 612 CTGCAATCAATATPACAACTATPACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 671
 606 CTGCAATCAATATPACAACTATPACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 665
 672 GTGTTGATGCTCTCATTTGAACTGCTGATGAAATCTGATGAAATCTGATGAAATCTGATG 731
 666 GTGTTGATGCTCTCATTTGAACTGCTGATGAAATCTGATGAAATCTGATGAAATCTGATG 725
 732 AGTTTCTCAAGTGCCTCAACCCCTCTTTCAACCCCTCTGAGAGAGAGAGAGAGAGAGAG 791
 726 AGTTTCTCAAGTGCCTCAACCCCTCTTTCAACCCCTCTGAGAGAGAGAGAGAGAGAGAG 785
 792 ATGAAAGTATGAGATGGAGCTGAGACCGAGTGGACTGTACCGCTGTGTGTGCTCT 851
 786 ATGAAAGTATGAGATGGAGCTGAGACCGAGTGGACTGTACCGCTGTGTGTGCTCT 845
 852 GTGAAATTTGGTCTGTACAGCCATGACCTGTGACCGAAAGATCAAGAGGGGGCCAGAG 911
 846 GTGAAATTTGGTCTGTACAGCCATGACCTGTGACCGAAAGATCAAGAGGGGGCCAGAG 905
 912 CCCAGAGAGAGAGAGATACAGAGATATGTCAGAGAGCTCCAAAGAGATCAGGAAGACAG 971
 906 CCCAGAGAGAGAGAGATACAGAGATATGTCAGAGAGCTCCAAAGAGATCAGGAAGACAG 965
 972 CTGAAAGAGACAGAGAGTGGACCAAGAGATCTAATGAGGAGGACAGACAGCTGTCT 1031
 966 CTGAAAGAGACAGAGAGTGGACCAAGAGATCTAATGAGGAGGACAGACAGCTGTCT 1025
 1032 TGGATCCAGATCTTCTCCACTTCCAGCTGAGTTCAGTATACAGAGTGTCTGTCTACA 1091
 1026 TGGATCCAGATCTTCTCCACTTCCAGCTGAGTTCAGTATACAGAGTGTCTGTCTACA 1085
 1092 GTGCCCAATCACCAGTATTTGCTTATATAGCAATGAGTTTATTTTGTATTTCTGTTT 1151
 1086 GTGCCCAATCACCAGTATTTGCTTATATAGCAATGAGTTTATTTTGTATTTCTGTTT 1145
 1152 GCAATAAAGATATGAAGTGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1211
 1146 GCAATAAAGATATGAAGTGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1205
 1212 GAGTGTCTTTAAGAGAACTGTAATGCTGCTGAGGCTGAGGCTAGTAAGAGAACTGC 1271
 1206 GAGTGTCTTTAAGAGAACTGTAATGCTGCTGAGGCTGAGGCTAGTAAGAGAACTGC 1265
 1272 ATCAGATTGAAG 1331
 1266 ATCAGATTGAAG 1325
 1332 CAGCAGCTGAGGAGTGCACAGATGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1391
 1326 CAGCAGCTGAGGAGTGCACAGATGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1385
 1392 AGAGTTTGGAGAGCTCCAGGAG 1451
 1386 AGAGTTTGGAGAGCTCCAGGAG 1445
 1452 GTACGATGACAGACCTCCAGATCCAGCATCTCTTGGTCCCAATTAAGCTCTCAGAT 1511

1446 GTCAGATGACAGACCTCCAGCATCCAGGCATCTCTGGTCCCAATTAAGCTCTCAGAT 1505
 1512 ACATAGCCCATCTGCTAGTTAAACCAGTGTCCCTCAGACTTGGATGGAGTTTCTGGGAGG 1571
 1506 ACATAGCCCATCTGCTAGTTAAACCAGTGTCCCTCAGACTTGGATGGAGTTTCTGGGAGG 1565
 1572 GTACACCCAAATGATGAGATGATCTGTATATCTTTGAGCCCTTAGCGACCTTAACCAAT 1631
 1566 GTACACCCAAATGATGAGATGATCTGTATATCTTTGAGCCCTTAGCGACCTTAACCAAT 1625
 1632 TTAAATAATACCTTTTACCAAGGTGCTATTTCTCTGTAAACAC- TTTTCTTGGCAAGT 1690
 1626 TTAAATAATACCTTTTACCAAGGTGCTATTTCTCTGTAAACAC- TTTTCTTGGCAAGT 1685
 1691 TGACTTATCTTCAATATTTATATATATATATATATATATATATATATATATATATTTTCTTGA 1750
 1686 TGACTTATCTTCAATATTTATATATATATATATATATATATATATATATATATTTTCTTGA 1745
 1751 CTAGG 1755
 1746 CTAGG 1750

RESULT 5
 AAI57940/c
 ID AAI57940 standard; cDNA; 2655 BP.
 XX
 AC AAI57940;
 XX
 DT 22-OCT-2001 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 143.
 XX
 KW Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200153312-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 26-DEC-2000; 2000WO-US034263.
 XX
 PR 23-DEC-1999; 98US-00471275.
 PR 21-JAN-2000; 2000US-00488725.
 PR 25-APR-2000; 2000US-00552317.
 PR 20-JUN-2000; 2000US-00598042.
 PR 19-JUL-2000; 2000US-00620312.
 PR 03-AUG-2000; 2000US-00653450.
 PR 14-SEP-2000; 2000US-00662191.
 PR 19-OCT-2000; 2000US-00693036.
 PR 29-NOV-2000; 2000US-00727344.
 XX
 (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
 PI Zhou P, Goodrich R, Drmanac RT;
 XX
 DR WPI; 2001-442253/47.
 DR P-PSDB; AAM38784.
 XX
 PT Novel nucleic acids and polypeptides, useful for treating disorders such
 PT as central nervous system injuries.
 XX
 PS Claim 1; SEQ ID NO 143; 10078pp; English.
 XX
 CC The invention relates to human nucleic acids (AAI57998-AAI61369) and the

AC AAL37538;
CX 08-JAN-2002 (first entry)
CX Human musculoskeletal system related polynucleotide SEQ ID NO 3903.
CX
CX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
CX anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; antitumor;
CX vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
CX cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
CX neurological disease; infection; human; secreted protein;
CX musculoskeletal system; ds.
CX Homo sapiens.
CX WO200155367-A1.
CX 02-AUG-2001.
CX 17-JAN-2001; 2001WO-US001338.
CX 31-JAN-2000; 2000US-0179065P.
CX 04-FEB-2000; 2000US-0180628P.
CX 24-FEB-2000; 2000US-0184664P.
CX 02-MAR-2000; 2000US-0186350P.
CX 16-MAR-2000; 2000US-0189874P.
CX 17-MAR-2000; 2000US-0190076P.
CX 18-APR-2000; 2000US-0198123P.
CX 19-MAY-2000; 2000US-0205515P.
CX 07-JUN-2000; 2000US-0209467P.
CX 28-JUN-2000; 2000US-0214886P.
CX 30-JUN-2000; 2000US-0215135P.
CX 07-JUL-2000; 2000US-0216647P.
CX 07-JUL-2000; 2000US-0216880P.
CX 11-JUL-2000; 2000US-0217487P.
CX 11-JUL-2000; 2000US-0217496P.
CX 14-JUL-2000; 2000US-0218290P.
CX 26-JUL-2000; 2000US-0220963P.
CX 26-JUL-2000; 2000US-0220964P.
CX 14-AUG-2000; 2000US-0224518P.
CX 14-AUG-2000; 2000US-0224519P.
CX 14-AUG-2000; 2000US-0225113P.
CX 14-AUG-2000; 2000US-0225214P.
CX 14-AUG-2000; 2000US-0225366P.
CX 14-AUG-2000; 2000US-0225267P.
CX 14-AUG-2000; 2000US-0225268P.
CX 14-AUG-2000; 2000US-0225370P.
CX 14-AUG-2000; 2000US-0225447P.
CX 14-AUG-2000; 2000US-0225757P.
CX 14-AUG-2000; 2000US-0225758P.
CX 14-AUG-2000; 2000US-0225759P.
CX 18-AUG-2000; 2000US-0226279P.
CX 22-AUG-2000; 2000US-0226681P.
CX 22-AUG-2000; 2000US-0226686P.
CX 22-AUG-2000; 2000US-0227182P.
CX 23-AUG-2000; 2000US-0227009P.
CX 30-AUG-2000; 2000US-0228924P.
CX 01-SEP-2000; 2000US-0229287P.
CX 01-SEP-2000; 2000US-0229343P.
CX 01-SEP-2000; 2000US-0229344P.
CX 01-SEP-2000; 2000US-0229345P.
CX 05-SEP-2000; 2000US-0229509P.
CX 05-SEP-2000; 2000US-0229513P.
CX 06-SEP-2000; 2000US-0230437P.
CX 06-SEP-2000; 2000US-0230438P.
CX 08-SEP-2000; 2000US-0231242P.
CX 08-SEP-2000; 2000US-0231243P.
CX 08-SEP-2000; 2000US-0231244P.
CX 08-SEP-2000; 2000US-0231413P.
CX 08-SEP-2000; 2000US-0231414P.
CX 08-SEP-2000; 2000US-0232080P.
CX 08-SEP-2000; 2000US-0232081P.
CX 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236377P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241211P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.

1441 ATGATTTTGGAAAGCAGGGTCCCTGCTGCTACAGCCAGTGACAGTCTGA 1500
3493 ATGCTTTTCTTACACTAGTATTTTAACTAGTCAAACTTCAAGAAAACAATTAACAA 3552
1501 ATGCTTTTCTTACACTAGTATTTTAACTAGTCAAACTTCAAGAAAACAATTAACAA 1560
3553 GTTCTCTGTCATATGTTTGTGAACCTTCTATTTTATTTAGTCTTCTATATTGCA 3612
1561 GTTCTCTGTCATATGTTTGTGAACCTTCTATTTTATTTAGTCTTCTATATTGCA 1620
3613 TTTAACTGTTTGTGAACCTTCTATTTTATTTAGTCTTCTATATTGCA 3672
1621 TTTAACTGTTTGTGAACCTTCTATTTTATTTAGTCTTCTATATTGCA 1680
3673 ATTAAGTGA 3682
1681 ATTAAGTGA 1690

RESULT 7
BX60526
ID AX60526 standard; cDNA; 1755 BP.
CX
CX BX60526;
CX
CX
CX 26-FEB-2003 (first entry)
CX cDNA encoding novel human musculoskeletal system antigen #2870.
CX Gene; ss; musculoskeletal system antigen; cancer; metastasis;
CX re-vascularisation; thrombosis; arteriosclerosis; mineral content;
CX cardiovascular condition; wound; injury; burn; angiogenesis; ulcer;
CX post-operative tissue repair; limb regeneration; neuronal growth;
CX neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
CX AIDS-related complex; chondrocyte growth; bone regeneration;
CX periodontal regeneration; tissue transport; bone graft; skin aging;
CX keratinocyte growth; hair loss; melanocyte growth; cell proliferation;
CX cell growth; organ transplant; cell differentiation; body height; weight;
CX hair colour; eye colour; skin; percentage of adipose tissue;
CX pigmentation; cosmetic surgery; metabolism; biorhythm; cardiac rhythm;
CX depression; tendency for violence; pain; reproductive capability;
CX hormone level; endocrine level; appetite; libido; memory; stress;
CX storage capability; fat content; lipid content; protein content;
CX carbohydrate content; vitamin content; cofactor content;
CX nutritional component.
CX
CX Homo sapiens.
CX
CX
CX US2002147140-A1.
CX
CX 10-OCT-2002.
CX
CX 17-JAN-2001; 2001US-00764877.
CX
CX 31-JAN-2000; 2000US-0179065P.
CX 04-FEB-2000; 2000US-0180628P.
CX 28-JUN-2000; 2000US-0214896P.
CX 07-JUL-2000; 2000US-0216647P.
CX 07-JUL-2000; 2000US-0216880P.
CX 11-JUL-2000; 2000US-0217487P.
CX 11-JUL-2000; 2000US-0217496P.
CX 14-JUL-2000; 2000US-0218220P.
CX 26-JUL-2000; 2000US-0220963P.
CX 26-JUL-2000; 2000US-0220964P.
CX 14-AUG-2000; 2000US-0224518P.
CX 14-AUG-2000; 2000US-0224519P.
CX 14-AUG-2000; 2000US-0225267P.
CX 14-AUG-2000; 2000US-0225268P.
CX 14-AUG-2000; 2000US-0225270P.
CX 14-AUG-2000; 2000US-0225447P.
CX 14-AUG-2000; 2000US-0225757P.
CX 14-AUG-2000; 2000US-0225758P.

22-AUG-2000; 2000US-0226868P.
30-AUG-2000; 2000US-0228924P.
01-SEP-2000; 2000US-0229287P.
01-SEP-2000; 2000US-0229343P.
01-SEP-2000; 2000US-0229344P.
01-SEP-2000; 2000US-0229345P.
05-SEP-2000; 2000US-0229509P.
05-SEP-2000; 2000US-0229513P.
08-SEP-2000; 2000US-0231413P.
21-SEP-2000; 2000US-0234223P.
21-SEP-2000; 2000US-0234274P.
25-SEP-2000; 2000US-0234997P.
27-SEP-2000; 2000US-0235834P.
29-SEP-2000; 2000US-0236327P.
29-SEP-2000; 2000US-0236367P.
29-SEP-2000; 2000US-0236382P.
29-SEP-2000; 2000US-0236389P.
29-SEP-2000; 2000US-0236370P.
02-OCT-2000; 2000US-0236802P.
02-OCT-2000; 2000US-0237037P.
02-OCT-2000; 2000US-0237038P.
02-OCT-2000; 2000US-0237039P.
13-OCT-2000; 2000US-0239352P.
20-OCT-2000; 2000US-0240960P.
20-OCT-2000; 2000US-0241785P.
20-OCT-2000; 2000US-0241809P.
01-NOV-2000; 2000US-0244617P.
17-NOV-2000; 2000US-0249299P.
08-DEC-2000; 2000US-0251856P.
08-DEC-2000; 2000US-0251868P.
08-DEC-2000; 2000US-0251869P.
(ROSE/) ROSEN C A.
(RUBE/) RUBEN S M.
(BARA/) BARASH S C.
Rosen CA, Ruben SM, Barash SC;
WFI; 2003-128199/12.
Isolated nucleic acid molecules encoding musculoskeletal system associated polypeptides, useful for detecting disorders, e.g. cancer.
Disclosure; SEQ ID NO 3903; 321pp; English.
The invention describes an isolated nucleic acid molecule comprising a sequence encoding musculoskeletal system associated polypeptides useful for detecting disorders, e.g., cancer or cancer metastases, in animals or humans. The nucleic acid: stimulates re-vascularisation of ischaemic tissues associated with conditions such as thrombosis, arteriosclerosis, and other cardiovascular conditions; treats wounds due to injuries, burns, post-operative tissue repair, and ulcers; stimulates angiogenesis and limb regeneration; stimulates neuronal growth; can treat and prevent neuronal damage occurring in certain disorders or neurodegenerative conditions, such as, Alzheimer's disease, Parkinson's disease, and AIDS-related complex; stimulates chondrocyte growth, thus they can be used to enhance bone and periodontal regeneration and aid in tissue transports or bone grafts; prevents skin aging due to sunburn by stimulating keratinocyte growth; prevents hair loss, since FGF family members activate hair-forming cells and promotes melanocyte growth; stimulates growth and differentiation of hematopoietic cells and bone marrow cells when used in combination with other cytokines; maintains organs before transplantation or for supporting cell culture of primary tissues; induces tissue of mesodermal origin to differentiate in early embryos; increases or decreases the differentiation or proliferation of embryonic stem cells, besides, haematopoietic lineage; modulates mammalian characteristics, such as, body height, weight, hair colour, eye colour, skin, percentage of adipose tissue, pigmentation, size, and shape (e.g., cosmetic surgery); modulates mammalian metabolism; changes mammal's metal state or physical state by influencing biorhythms; cardiac rhythms, depression, tendency for violence, tolerance for pain, reproductive capabilities, hormonal or endocrine levels, appetite, libido, memory, or

CC stress; increases or decreases storage capabilities, fat content, lipid,
 CC protein, carbohydrate, vitamins, minerals, cofactors or other nutritional
 CC components. This sequence encodes a novel human musculoskeletal system
 CC antigen. Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from the US patent office at
 CC ftp.segdata.uspto.gov/sequence.html?DocID=20020147140
 XX
 XX Sequence 1755 BP; 500 A; 356 C; 351 G; 546 T; 0 U; 2 Other;

Query Match 45.4%; Score 1678.4; DB 7; Length 1755;
 Best Local Similarity 99.5%; Pred. No. 0;
 Matches 1682; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

2y 1993 AACATTAATTTGTTTATTTATTTCTGAGAAGTTGAGGCAATTCAGTATTCCTCCAGGAT 2052
 Db 1 ANANNTTTTGTGTTTTTATTTATTTCTGAGAAGTTGAGGCAATTCAGTATTCCTCCAGGAT 60
 2y 2053 GGCACAAAGGCGACCAAGCAGGCTTAGGATATCCAGCCTACCAATATGCTCATTCGA 2112
 Db 61 GGCACAAAGGCGACCAAGCAGGCTTAGGATATCCAGCCTACCAATATGCTCATTCGA 120
 2y 2113 CTAAGTGGAGGGTGAAGTGGCCCTGCTCTCTTTTCTGGACCTCAGTTCCCTCAGT 2172
 Db 121 CTAAGTGGAGGGTGAAGTGGCCCTGCTCTCTTTTCTGGACCTCAGTTCCCTCAGT 180
 2y 2173 GAGCTGTAAGAATGCACATAACCTTTTGAATTTGATAGTTTATAAATCTGTGTTCTGAT 2232
 Db 181 GAGCTGTAAGAATGCACATAACCTTTTGAATTTGATAGTTTATAAATCTGTGTTCTGAT 240
 2y 2233 CATTTGTCGAGGGGAGATAGTTTCTGTGATTTTCTTCTCTATAGATAAATG 2292
 Db 241 CATTTGTCGAGGGGAGATAGTTTCTGTGATTTTCTTCTCTATAGATAAATG 300
 2y 2293 AAATCTTTGTTACTAGAACAAAGAAATGTCAGATGGCCAAACAAAGATGACCAATTTGAT 2352
 Db 301 AAATCTTTGTTACTAGAACAAAGAAATGTCAGATGGCCAAACAAAGATGACCAATTTGAT 360
 2y 2353 CTCAGCTGATGACCTTACAGGCTGCTATGATAGGATCTCTATGGTAAAGCAGA 2412
 Db 361 CTCAGCTGATGACCTTACAGGCTGCTATGATAGGATCTCTATGGTAAAGCAGA 420
 2y 2413 AGAGTGGGGAAGAGAACCAACCCACTCTGTCTCATATTTGCAATTTCACTTAACT 2472
 Db 421 AGAGTGGGGAAGAGAACCAACCCACTCTGTCTCATATTTGCAATTTCACTTAACT 480
 2y 2473 CCGCTCGAAATAGAAAGCATTCCTTAGAGATGAGGATAAAGAAAGTTTCAGATTCAA 2532
 Db 481 CCGCTCGAAATAGAAAGCATTCCTTAGAGATGAGGATAAAGAAAGTTTCAGATTCAA 540
 2y 2533 CAGGGGAAGAAATGAGATTTAATCCTAAACTGTGACTTGGGAGGTCAGTCATTTA 2592
 Db 541 CAGGGGAAGAAATGAGATTTAATCCTAAACTGTGACTTGGGAGGTCAGTCATTTA 600
 2y 2593 CAGTTAGCTGTCTTTCGACTTCTGTGATTTAAACCCACTCACTACCTGTTTCA 2652
 Db 601 CAGTTAGCTGTCTTTCGACTTCTGTGATTTAAACCCACTCACTACCTGTTTCA 660
 2y 2653 GATGCTTTGGAATACCAAGATTAATCTTGACATAAGATCTCTTTGCAAGAGAGCAG 2712
 Db 561 GATGCTTTGGAATACCAAGATTAATCTTGACATAAGATCTCTTTGCAAGAGAGCAG 720
 2y 2713 ATTAAGACCATCAGAAGAAATTTATTAGTTGTTAATGACAGGCAACTGTGAGAACT 2772
 Db 721 ATTAAGACCATCAGAAGAAATTTATTAGTTGTTAATGACAGGCAACTGTGAGAACT 780
 2y 2773 GTTCTGCGAAATAGAAATCTCTAGTTTCTTCTGTTCTCTATTTGAAGGAGAAAT 2832
 Db 781 GTTCTGCGAAATAGAAATCTCTCTAGTTTCTTCTGTTCTCTATTTGAAGGAGAAAT 840
 2y 2833 CCACCTTTGTTAGCATTTCAAGCTTTTATGATCCATCCCATCTAAACACTCTTCAACT 2892
 Db 841 CCACCTTTGTTAGCATTTCAAGCTTTTATGATCCATCCCATCTAAACACTCTTCAACT 900

Qy 2893 CCACCTTTGTTAGTCTGAAATGCAAGCTCCTCTCTCAAGTGCCTTGGAGAACTCAGCAGC 2952
 Db 901 CCACCTTTGTTAGTCTGAAATGCAAGCTCCTCTCTCAAGTGCCTTGGAGAACTCAGCAGC 960
 Qy 2953 AGCCCTTAATCAAGGTTTTTACCAAGCCCTTGGACACTATGGAGAGGAGGCAAGAGTACAC 3012
 Db 961 AGCCCTTAATCAAGGTTTTTACCAAGCCCTTGGACACTATGGAGAGGAGGCAAGAGTACAC 1020
 Qy 3013 CAATTTGTTAAAGCAAGAAACCAAGCTGCTCTCTCACTAGTCAATTTAGAACATGGTTAT 3072
 Db 1021 CAATTTGTTAAAGCAAGAAACCAAGCTGCTCTCTCACTAGTCAATTTAGAACATGGTTAT 1080
 Qy 3073 CATCCAGACTACTCTACCTTGCAACATTGAATCCCAAGAGCAAAATCCCAATTCCTCTT 3132
 Db 1081 CATCCAGACTACTCTACCTTGCAACATTGAATCCCAAGAGCAAAATCCCAATTCCTCTT 1140
 Qy 3133 GAGTCTCCAGCTTCTGTGTAATAGGAGAGCTGTCTGTATGCGGTAGATAATCATGAT 3192
 Db 1141 GAGTCTCCAGCTTCTGTGTAATAGGAGAGCTGTCTGTATGCGGTAGATAATCATGAT 1200
 Qy 3193 CTGAGGACCATTCATGGAAGCTGCTAAATAGCTAGTCTGGGAGTCTTCCATAAAGTTT 3252
 Db 1201 CTGAGGACCATTCATGGAAGCTGCTAAATAGCTAGTCTGGGAGTCTTCCATAAAGTTT 1260
 Qy 3253 TGCATGGAGCAAAACAAACAGGATTAATAGTTTGGTTTCTTCAGCCCTCTAAAGCAT 3312
 Db 1261 TGCATGGAGCAAAACAAACAGGATTAATAGTTTGGTTTCTTCAGCCCTCTAAAGCAT 1320
 Qy 3313 AGGCTTAGCTGCTGAGGCTTCTTGGGCTTCTCTGTGTGTGTAGTTTGTGTAACACTAT 3372
 Db 1321 AGGCTTAGCTGCTGAGGCTTCTTGGGCTTCTCTGTGTGTGTAGTTTGTGTAACACTAT 1380
 Qy 3373 AGCATCTGTTAAGATCCAGTGTCCATGGAACCTTCCACATGCGGTGACTCTCGACTAT 3432
 Db 1381 AGCATCTGTTAAGATCCAGTGTCCATGGAACCTTCCACATGCGGTGACTCTCGACTAT 1440
 Qy 3433 ATCAGTTTTTGGAAAGCAGGTTCTCTGCTGCTAAACAGCCACCTGGACCACTGTA 3492
 Db 1441 ATCAGTTTTTGGAAAGCAGGTTCTCTGCTGCTAAACAGCCACCTGGACCACTGTA 1500
 Qy 3493 ATGCTTTTCTTTTACACCTATGTTTTTAAAGTAGTCAAACTTCAAGAAACAATCTAAACAA 3552
 Db 1501 ATGCTTTTCTTTTACACCTATGTTTTTAAAGTAGTCAAACTTCAAGAAACAATCTAAACAA 1560
 Qy 3553 GTTCTGTGCTATGTTTGTGAACCTTGTATTTTGTATTTAGTAGGCTTCTATATTGCA 3612
 Db 1561 GTTCTGTGCTATGTTTGTGAACCTTGTATTTTGTATTTAGTAGGCTTCTATATTGCA 1620
 Qy 3613 TTTAACTTTGTTTTTGTAACTCCTGATTCCTTTTCGGATCTATTGATGAATAAAGAA 3672
 Db 1621 TTTAACTTTGTTTTTGTAACTCCTGATTCCTTTTCGGATCTATTGATGAATAAAGAA 1680
 Qy 3673 ATTAAGTGA 3682
 Db 1681 ATTAAGTGA 1690

RESULT 8
 ABO88210
 ID ABO88210 standard; cDNA; 1010 BP.
 XX
 AC ABO88210;
 XX
 DT 18-SEP-2002 (first entry)
 XX
 DE Human osteoblast differentiation related cDNA SEQ ID NO 117.
 XX
 KW Human; osteoblast; stem cell differentiation; bone tissue deposition;
 XX osteoporosis; osteopathic; ss.
 OS Homo sapiens.
 XX

PN WO200250301-A2.
 XX 27-JUN-2002.
 PD 18-DEC-2001; 2001WO-US048276.
 PF 18-DEC-2000; 2000US-0255882P.
 PR 24-APR-2001; 2001US-0285691P.
 XX (GENE-) GENE LOGIC INC.
 PA (PROC) PROCTER & GAMBLE CO.
 XX Ji D, Axelrod DW, Cook JS, Jaiswal N, Binstein R, Roughton A;
 PI Mertz L;
 XX WPI; 2002-557663/59.
 DR Use of genes and their expression profiles associated with osteoblast
 PT differentiation for screening modulators bone formation, for diagnosing
 PT or treating e.g. osteoporosis, or as markers for the differentiation
 PT process.
 XX Claim 1; SEQ ID NO 117; 78pp + Sequence Listing; English.
 PS The invention relates to genes and their expression profiles are used
 CC for: (a) screening modulators of precursor stem cell differentiation into
 CC osteoblasts, or bone tissue deposition; (b) diagnosing abnormal
 CC osteoblasts of bone tissue, abnormal rate of osteoblast formation or
 CC osteoporosis; or (c) treating or monitoring treatment of the conditions
 CC cited in (b), or monitoring the progression of bone tissue deposition.
 CC Specific conditions include postmenopausal osteoporosis, glucocorticoid
 CC osteoporosis or male osteoporosis, osteopenia, osteodystrophy, drug-
 CC induced abnormalities in bone formation or bone loss, conditions that
 CC involve altered bone metabolism (e.g. idiopathic juvenile osteoporosis),
 CC skeletal disease linked to breast cancer, mastocytosis, Fanconi syndrome
 CC or fibrous dysplasia. The present sequence is that of an osteoblast
 CC differentiation associated cDNA marker of the invention. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 1010 BP; 295 A; 211 C; 190 G; 314 T; 0 U; 0 Other;

Query Match 27.3%; Score 1008.4; DB 6; Length 1010;
 Best Local Similarity 99.9%; Pred. No. 2.28-248;
 Matches 1009; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 2683 TTGACATAAGATCTCATTTGAGAAAGCAGATTAAAGACCATCAGAGGAAATTTATTAG 2742
 1 TTGACATAAGATCTCATTTGAGAAAGCAGATTAAAGACCATCAGAGGAAATTTATTAG 60
 2743 GTTGTAAATGACAGCAACTGTGAAACTGTTGTGCAAAATAGAAATCTCTTCTAGTT 2802
 61 GTTGTAAATGACAGCAACTGTGAAACTGTTGTGCAAAATAGAAATCTCTTCTAGTT 120
 2803 TTTCTTGTCTCATTTGAAAGGAGAAATCCACTTTGTTAGCATTTCAAGCTTTATG 2862
 121 TTTCTTGTCTCATTTGAAAGGAGAAATCCACTTTGTTAGCATTTCAAGCTTTATG 180
 2863 TATCCATCCCATCTATAAACTCTTCAAACTCCACTTGTTCAGTCTGAAATGCACTCCCT 2922
 181 TATCCATCCCATCTATAAACTCTTCAAACTCCACTTGTTCAGTCTGAAATGCACTCCCT 240
 2923 GTCCAGTCCCTTGGAGAACTCAGAGCAAGCGCTTAATCAAGGTTTACAGCCCTT 2982
 241 GTCCAGTCCCTTGGAGAACTCAGAGCAAGCGCTTAATCAAGGTTTACAGCCCTT 300
 2983 GGACACTATGGAGAGGAGGAGAGTACACCAATTTGTTAAAGCAAGAAACACACAGTGT 3042
 301 GGACACTATGGAGAGGAGGAGAGTACACCAATTTGTTAAAGCAAGAAACACACAGTGT 360
 3043 CTCCTCAGTCTATTAGACATGGTTATCATCCAGACTCTACTCCCTGCAACATTG 3102

Db 361 CTCCTCAGTCTATTAGACATGGTTATCATCCAGACTCTACTCCCTGCAACATTG 420
 QY 3103 AACTCCCAAGAGCAAAATCCACATTCCTCTTGAGTTCTGAGCTTCTGTGTAATAAGGCA 3162
 Db 421 AACTCCCAAGAGCAAAATCCACATTCCTCTTGAGTTCTGAGCTTCTGTGTAATAAGGCA 480
 QY 3163 GCTGCTGCTATGCGGTAGATACATGATCTGAGGACCATTCATGGAAGCTGCTAAATA 3222
 Db 481 GCTGCTGCTATGCGGTAGATACATGATCTGAGGACCATTCATGGAAGCTGCTAAATA 540
 QY 3223 GCTAGTCTGCGGAGTCTTCCATAAAGTTTTCATGAGCAAAACAAACAGGATTAACCTA 3282
 Db 541 GCTAGTCTGCGGAGTCTTCCATAAAGTTTTCATGAGCAAAACAAACAGGATTAACCTA 600
 QY 3283 GCTTTGGTTCTTTCAGCCCTCTAAAGCATAGGCTTAGCCCTGACGGCTTCTCTGGGCTT 3342
 Db 601 GCTTTGGTTCTTTCAGCCCTCTAAAGCATAGGCTTAGCCCTGACGGCTTCTCTGGGCTT 660
 QY 3343 TCTCTGCTGCTAGTCTTTTAAACACATATAGCATCTGTTAAGATCCAGTGTCCATGGAA 3402
 Db 661 TCTCTGCTGCTAGTCTTTTAAACACATATAGCATCTGTTAAGATCCAGTGTCCATGGAA 720
 QY 3403 ACCTTCCCATGCGGTGACTCTGACTATATCAGTCTTTTGGAAAGCAGGTTCTCTCTGC 3462
 Db 721 ACATTTCCCATGCGGTGACTCTGACTATATCAGTCTTTTGGAAAGCAGGTTCTCTCTGC 780
 QY 3463 CTGCTAAACAAGCCACGCTGGACAGCTCTGAATGCTTTTCCCTTACACCTATGTTTTTAAG 3522
 Db 781 CTGCTAAACAAGCCACGCTGGACAGCTCTGAATGCTTTTCCCTTACACCTATGTTTTTAAG 840
 QY 3523 TAGTCAAACTTCAAGAAACAATCTAAACAAGTTTCTGTTGCATATGTTGTTGTAAGCTTG 3582
 Db 841 TAGTCAAACTTCAAGAAACAATCTAAACAAGTTTCTGTTGCATATGTTGTTGTAAGCTTG 900
 QY 3583 TATTTGTTATTTAGTGGCTCTCTATATTCATTTAACTTTGTTTCTAACTCTGATTTCTT 3642
 Db 901 TATTTGTTATTTAGTGGCTCTCTATATTCATTTAACTTTGTTTCTAACTCTGATTTCTT 960
 QY 3643 CCTTTTCGATCTATTGATGAATAAGAAATTAAGTGCAAAAAA 3692
 Db 961 CCTTTTCGATCTATTGATGAATAAGAAATTAAGTGCAAAAAA 1010

RESULT 9
 AAX39840
 ID AAX39840 standard; DNA; 1017 BP.
 XX
 AC AAX39840;
 XX
 DT 02-JUL-1999 (first entry)
 XX
 DE Gastric cancer associated gene.
 XX
 KW Cancer associated antigen; diagnosis; research; treatment; human;
 KW breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;
 KW prostate cancer; ss.
 XX
 OS Homo sapiens.
 XX
 PN MO9904265-A2.
 XX
 PD 28-JAN-1999.
 XX
 PF 15-JUL-1998; 98WO-US014679.
 XX
 PR 17-JUL-1997; 97US-00896164.
 PR 10-OCT-1997; 97US-0061599P.
 PR 10-OCT-1997; 97US-0061765P.
 PR 10-OCT-1997; 97US-00948705.
 PR 11-OCT-1997; 97GB-00021697.
 PR 22-JUN-1998; 98US-00102322.
 XX
 PA (LUDW-) LUDWIG INST CANCER RES.

```
Db 618 TCGCCCTAGATACAGAGTCTTATACCCACACACCTGTCTACAGTGTCAATCAATAAGTG 677
Qy 806 -ACGTCCTTGTGAAAAAACAATAAACC 837
Db 678 CACGTCCTTGTGAAAAAACAATAAACC 710

RESULT 15
CUI75286
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

CUI75286 712 bp mRNA linear EST 19-MAY-2003
AGENCOURT_13978429 NIH_MGC_172 Homo sapiens cDNA 5', mRNA sequence.
CUI75286
CUI75286.1 GI:30859713
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaaps-remail.nih.gov
Tissue Procurement: Dr. Jamie Thompson, University of WI
cDNA Library Preparation: Gina Zastrow-Hayes
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ILLNL at:
http://image.llnl.gov
Plate: NDKM49 row: m column: 06
High quality sequence start: 31
High quality sequence stop: 712.
Location/Qualifiers
1..712
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, WA01, passage 39"
/lab_host="DH10B TonA"
/clone_lib="NIH MGC 172"
/notes="Vector: pDONR201; Site 1: attP2; Site 2: attP1;
LIBR_PRIMING - oligo dT; METHOD - full-length enriched;
Embryonic Stem Cells HI; LIBR PROVIDER - Bradfield"

FEATURES
Source
Query Match 60.1%; Score 666.4; DB 14; Length 712;
Best Local Similarity 98.3%; Pred. No. 9.6e-125;
Matches 684; Conservative 0; Mismatches 11; Indels 1; Gaps 1;

Qy 143 GAGAACTGAACGACAGGGGAAAGAGTCTCAGTACGACCGTCCCGAGCATCCGGACA 202
Db 17 GAGTAAAAAAGACGACGGGAAAGAGGTCTCAGTACGACCGTCCCGAGCATCCGGACA 76
Qy 203 CCACAGCGCCCTTCGCTCCAGCGAGAAACACACTTCTCAAACTTCACCTCAACACTT 262
Db 77 CCACAGCGCCCTTCGCTCCAGCGAGAAACACACTTCTCAAACTTCACCTCAACACTT 136
Qy 263 CTTTCCCAAGCCAGAGATGCAAGAGGAAACATGAGTGGTGTGCTGGGGGACC 322
Db 137 CTTTCCCAAGCCAGAGATGCAAGAGGAAACATGAGTGGTGTGCTGGGGGACC 196
Qy 323 CCCCAGCAGCATCTTCCAGGTCCACCGTGATCAACATCCACGAGACCTCCGTGCC 382
Db 197 CCCCAGCAGCATCTTCCAGGTCCACCGTGATCAACATCCACGAGACCTCCGTGCC 256
Qy 383 CGACCATGTGCTGTGCTCCCTGTTCACACCCCTCTTCTTGAACGTGCTGTCTGGGCTT 442
Db 257 CGACCATGTGCTGTGCTCCCTGTTCACACCCCTCTTCTTGAACGTGCTGTCTGGGCTT 316
Qy 443 CATAGCAATCGCCTACTCCGTGAAGTCTAGGACAGGAAGATGGTTGGCGACGTGACCGG 502
```

```
Db 317 CATAGCAATCGCCTACTCCGTGAAGTCTAGGACAGGAAGATGGTTGGCGACGTGACCGG 376
Qy 503 GGCCCAAGCCCTATGCTCCACCCGCAAGTGCCTGAACATCTGGGCCCTGATTTGGGCAT 562
Db 377 GGCCCAAGCCCTATGCTCCACCCGCAAGTGCCTGAACATCTGGGCCCTGATTTGGGCAT 436
Qy 563 CCTCATGACCAATGGATTTCATCCTGTTACTTGGTATTTCGGCTCTGTGACAGTCTACCATAT 622
Db 437 CCTCATGACCAATGGATTTCATCCTGTTACTTGGTATTTCGGCTCTGTGACAGTCTACCATAT 496
Qy 623 TATGTTACAGATAATACAGGAAAAAGGGGTTACTAGTAGCCGCCCATAGCCTGCAACCT 682
Db 497 TATGTTACAGATAATACAGGAAAAAGGGGTTACTAGTAGCCGCCCATAGCCTGCAACCT 556
Qy 683 TTGCACCTCCACTGTGCAATGCTGGCCCTGCACGCTGGGGCTGTGTGCCCTGCCCCCTTGG 742
Db 557 TTGCACCTCCACTGTGCAATGCTGGCCCTGCACACTGGGGCTGTGTGCCCTGCCCCCTTGG 616
Qy 743 TCTTGGCCCTTAGATACAGCAGTATTATACCCACACACCTGTCTACACTGACATTCAATAAA 802
Db 617 TCTTGGCCCTTAGATACAGCAGTATTATACCCACACACCTGTCTACAGTGTCAITCAATAAA 676
Qy 803 GTG-ACGTGCTTGTGAAAAAACAATAAACC 837
Db 677 GTGACGCTGCTTGTGAAAAAACAATAAACC 712
```

Search completed: March 11, 2004, 14:23:57
Job time : 1840.29 secs


```

      840 GAGGAGATGACCAAGATATGTCAGGAGCTCCAAAGAGCATCAGGAAACAGCTGAAAGACC 859
      QY      983 AAGAGAGTGAGCACCAAAGAGATCTAA 1009
      Db      900 AAGAGAGTGAGCACCAAAGAGATCTAA 926

RESULT 11
ADB58526
ID ADB58526 standard; DNA; 1370 BP.

```

Query Match 24.7%: Score 911.8: DB 2: Length 926:

OV 83 ATGTGGAAACGCTGGCTCGCGCTCGCGCTGGTGGCGGTGCCCTGGGTCCGCGCC 142

db
1 ATGTGAAACGCTGGCTCGCGCTCGCGCTGGTGGCGGTCCCTGGGTCCGCGCC 60

QY 143 GAGGAAGAGCTAAGGAGCAATCCAAGATCTGTGCCAATGTGTTTGTGGAGCCGGCCGG 202

[illegible]

Db 121 GAATGTCAGTCA CAGAGAAAGGGAA CCACTGTCTCTGCA TTGAGCAATGCAAACCT 180

[illegible]

323 CGAGATGCCTGCCTCACTGGATCCAAAATCCAGGTTGATTACGATGGACA CTGCAAAAGAG 382

Db 241 CGAGATGCCCTGCCCTCACTGGATCCAAATCCAGGTTGATTACGATGGACACTGCAAGAG 300

QY 443 CTCCGACGTCGCATCATCCAGTGGCTGGAAGCTGAGATCATCTCAGATGGCTGGTTCTCT 502

DB 361 CTCGACGTCGCATCATCCAGTGGCTGGAAAGCTGAAGATCATTCACAGATGGCTGGTCTCT 420

190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000


Qy 563 TCTCGCCTGGACTCCAGTGAATTCCTGAAGTTTGTGGAACAGAAATGAAACTGCCATCAAT 622

22

Db 540 ATTACAAGTATCCAGACCAGGAGAACAAAGTTGCTTAGGGGACTCTGTGTTGATGCT 599

[illegible][illegible]

D**b** 660 TGCC TCAACCCATCTTTCAACCCCTCCTGAGAAAGTGTGCCCTGGAGGATGAAACGTAT 719



863 GTCTGTACAGCCATGACCTGTGACGGAAGAATCAGAAAGGGGCCAGACCAGACAGAG 922

DB 780 GTCTGTACAGCCATGACCTGTGACGGGAAGAAATCAGAAAGGGGGCCAGACCCAGACACAGAG 839

553 ATAATGGTGATTCTCGCTGGCACTCCAGTGAATTCCTGGAAGTTGTGGACACAATGAAA 613
 612 CTGCCATAAATATTACAACTGATCCAGACCAGGAGAACAAACAAGTTGCTTTAGGGACTCT 671
 613 CTGCCATAAATATTACAACTGATCCAGACCAGGAGAACAAACAAGTTGCTTTAGGGACTCT 672
 672 GTGTTGATGCTCTCATTTGAACTGTCTGATGAAAAATGCTGATTGGAACACTCAGCTTCCAG 731
 673 GGTGTTGATGCTCTCATTTGAACTGTCTGATGAAAAATGCTGATTGGAACACTTAGCTTTCAAG 732
 732 -AGTTTTCTCAAAGTGCC-TCAAACCCATCTTTCAACCCCTCTCGAGAAGAGTGTGCCCTGGA 789
 733 AGTTTTCTCAAAGNGCCTTNAACCCATCTTTNAACTTTCTTGAGAAGAAATGTGSCCTTGGA 792
 790 GGATGAAACGATG-CAGATGAGC-TGAGACCGAGTGAGCTGTAAACGCCTGTGTCTGT 847
 793 GGATGAAACGATGTGCANATGAGCTTGAANCGAGTGGACTGTAAACCGTTGGNCTGG 852
 848 GCCTGTGGAAA 858
 853 GNCCTGGGGAA 863

RESULT 14
 ABZ34644
 ID ABZ34644 standard; cDNA; 930 BP.
 AC ABZ34644;
 XX
 XX DT 04-FEB-2003 (first entry)
 XX
 DE Coding sequence SEQ ID 2, upregulated in osteogenesis.
 XX
 KW Osteopathic; osteogenesis modulator; gene therapy; osteogenesis;
 XX Osteoporosis; bone disease; downregulator; human; ss.
 XX Homo sapiens.
 OS
 WO200281745-A2.
 PN
 PD 17-OCT-2002.
 XX
 PF 05-APR-2002; 2002WO-IB002211.
 PR
 PP 05-APR-2001; 2001US-0281400P.
 XX
 PA (AVET) AVENTIS PHARMA SA.
 PI Garcia T, Roman Roman S, Baron R, Call K, Theilhaber J;
 PI Connolly T, Jackson A, Bushnell SE, Rawadi G;
 XX WPI; 2003-058567/05.
 DR
 PT Novel isolated nucleic acid upregulated/downregulated in osteogenesis,
 PT useful for bone disease therapy in subject.
 XX
 PS Claim 26; Page 84; 237pp; English.
 XX
 CC The present invention relates to novel nucleotide sequences, which are
 CC differentially expressed in models of osteogenesis upon being put in
 CC contact with a stimulator of osteogenesis. The present sequence is one
 CC such sequence. This sequence can be used for diagnosing osteoporosis/bone
 CC disease in a patient, promoting osteogenesis and/or preventing
 CC osteoporosis/bone disease. The present sequence encodes a secreted
 CC protein
 XX
 SQ Sequence 930 BP; 263 A; 236 C; 247 G; 184 T; 0 U; 0 Other;

Query Match 19.8%; Score 729.4; DB 7; Length 930;
 Best Local Similarity 87.8%; Pred. No. 1.1e-176;
 Matches 796; Conservative 0; Mismatches 111; Indels 0; Gaps 0

Oy 103 GCTCGCCTCGCCTGGTGGCGGTCCCTCGCGCCGAGGAGAGCTTAAGGACAA 162

X Cancer associated antigen; diagnosis; research; treatment; human;
W Breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;
M prostate cancer; ss.
X Homo sapiens.
S WO9904265-A2.
N 28-JAN-1999.
X 15-JUL-1998; 98WO-US014679.
R 17-JUL-1997; 97US-00896164.
R 10-OCT-1997; 97US-0061599P.
R 10-OCT-1997; 97US-0061765P.
R 10-OCT-1997; 97US-00948705.
R 11-OCT-1997; 97GB-00021697.
R 22-JUN-1998; 98US-00102322.
X (LUDW-) LUDWIG INST CANCER RES.
X Old LJ, Scanlan MJ, Stockert E, Gure A, Chen Y, Gout I;
X O'hare M, Obata Y, Pfreundschuh M, Tureci O, Sahin U;
X WPI; 1999-132448/11.
X New isolated cancer associated nucleic acids and polypeptides - isolated
X using sera from cancer patients, used to develop products for the
X diagnosis, monitoring or treatment of cancers.
X Claim 67; Page 568; 787pp; English.
X The invention relates to a method for diagnosing a disorder characterised
X by expression of a human cancer associated antigen precursor coded for by
X a nucleic acid molecule (NAM). The method comprises: (a) contacting a
X biological sample isolated from a subject with an agent that specifically
X binds to the NAM, an expression product or a fragment of an expression
X product complexed with an HLA molecule; and (b) determining the
X interaction between the agent and the NAM or the expression product as a
X determination of the disorder. The products and methods can be used in
X the diagnosis, monitoring, research, or treatment of conditions
X characterised by the expression of various cancer associated antigens.
X The invention provides nucleic acid sequences and encoded polypeptides
X which are cancer associated antigen precursors expressed in human breast
X cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and
X lung cancer
X Sequence 810 BP; 234 A; 186 C; 206 G; 177 T; 0 U; 7 Other;
X
X Query Match 19.7%; Score 729.2; DB 2; Length 810;
X Best Local Similarity 96.4%; Pred. No. 1.2e-176;
X Matches 784; Conservative 0; Mismatches 25; Indels 4; Gaps 4;
X
X 207 GTCCAGTCCAGAGAAGGGGAACCCACCTGCTCTGCAATTGAGCAATCCAAACCTCACA 266
X 1 GTCCAGTCCAGAGAAGGGGAACCCACCTGCTCTGCAATTGAGCAATCCAAACCTCACA 60
X
X 267 AGAGGCGCTGTGTGCGAGTAATGGCAAGACCTTACCTCAACCACTGTGAACCTGATCGAG 326
X 61 AGAGGCGCTGTGTGCGAGTAATGGCAAGACCTTACCTCAACCACTGTGAACCTGATCGAG 120
X
X 327 ATCCCTGCCTCACTGGATCCAAATCCAGTTGATTAACGATGACACATGCCAAAGAGAAGA 386
X 121 ATCCCTGCCTCACTGGATCCAAATCCAGTTGATTAACGATGACACATGCCAAAGAGAAGA 180
X
X 387 AATCCGTAGTCCATCTGCCAGCCCAAGTTGTTGCTATCAGTCCACCGTGATGAGCTCC 446
X 181 AATCCGTAGTCCATCTGCCAGCCCAAGTTGTTGCTATCAGTCCACCGTGATGAGCTCC 240
X
X 447 GACGTGCGATCATCCAGTGGCTGGAGCTGAGATCATCTCAGATGGCTGTTCTCTAAAG 506
X 241 GACGTGCGATCATCCAGTGGCTGGAGCTGAGATCATCTCAGATGGCTGTTCTCTAAAG 300

QY 507 GCAGCACTACAGTGAATCCTAGACAGAGTATTTTAAGAACTTTGATATGATGATTC 566
Db 301 GCAGCACTACAGTGAATCCTAGACAGAGTATTTTAAGAACTTTGATATGATGATTC 360
QY 567 GCCTGGACTCCAGTGAATCCTGAAAGTTTGTGGACAGAACTGAACTGCCATCAATATTA 626
Db 361 GCCTGGACTCCAGTGAATCCTGAAAGTTTGTGGACAGAACTGAACTGCCATCAATATTA 420
QY 627 CAACGTATCCAGACCCAGGAGAACAACTAGTTGCTTAGGGGACTCTGTTGATGCTCTCA 686
Db 421 CAACGTATCCAGACCCAGGAGAACAACTAGTTGCTTAGGGGACTCTGTTGATGCTCTCA 480
QY 687 TTGAACCTGTCTGATGAAATGCTGATTTGAAATCTCAGCTTCCAGAGTCTTCTCAAGTGCC 746
Db 481 TTGAACCTGTCTGATGAAATGCTGATTTGAAATCTCAGCTTCCAGAGTCTTCTCAAGTGCC 540
QY 747 TCAACCCATCTTTCAACCCCTCCTGAGAAAGTGTGCCCTGGAGGATGAAACGTATGCGAG 806
Db 541 TCAACCCATCTTTCAACCCCTCCTGAGAAAGTGTGCCCTGGAGGATGAAACGTATGCGAG 600
QY 807 ATGGAGCTGAGACCCGAGGTGGACTGTAA - CCGCTGTGTCTGTGCTGTGAAATTTGGGTC 865
Db 601 ATGGAGCTGAGACCCGAGGTGGACTGTAA - CCGCTGTGTCTGTGCTGTGAAATTTGGGTC 660
QY 866 TGTACAGCCATGACCTGTGTGACGGAAAGAAATCAGAAAGGGGCCCCAGACCCAGAGAGGAG 925
Db 661 TGT - CAGCCATGACCTGTGACNGAAAGAAATCAGAAAGGGGCCCCAGACCCAGAGAGGAG 719
QY 926 GAGATGACCCAGATATGTCACAGGAGCTCCAAAGACATCAGGAAACAGCTGAAAGACCAAG 985
Db 720 GAGATGACCCAGATATGTCACAGGAGCTCC - AAAGCTTAGGAAACAGCTTGAAGAACCCAG 778
QY 986 AGAGTGAAGCCAAAGAGATCTTAATGAGGAGGC 1018
Db 779 AGAGGAGGCCCCAA - AGAGATTAATGAGGAGGC 810

Search completed: March 10, 2004, 10:59:23
Job time : 904.804 secs

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	3693	100.0	4	3693	4	US-09-976-594-908	Sequence 908, Appl
2	892.8	24.2	896	4	US-09-621-976-15155	Sequence 15155, A	Sequence 209, Appl
3	619.4	16.8	621	4	US-09-643-597-209	Sequence 209, Appl	Sequence 209, Appl
4	619.4	16.8	621	4	US-09-480-884A-209	Sequence 209, Appl	Sequence 209, Appl
5	619.4	16.8	621	4	US-09-542-615A-209	Sequence 209, Appl	Sequence 209, Appl
6	619.4	16.8	621	4	US-09-606-421B-209	Sequence 209, Appl	Sequence 209, Appl
7	477.2	12.9	708	3	US-08-896-164-34	Sequence 34, Appl	Sequence 315, Appl
8	336	9.1	336	4	US-09-643-597-315	Sequence 315, Appl	Sequence 315, Appl
9	336	9.1	336	4	US-09-480-884A-315	Sequence 315, Appl	Sequence 315, Appl
10	336	9.1	336	4	US-09-542-615A-315	Sequence 315, Appl	Sequence 315, Appl
11	336	9.1	336	4	US-09-606-421B-315	Sequence 315, Appl	Sequence 315, Appl
12	70.2	1.9	1355	4	US-09-412-554A-7	Sequence 7, Appl	Sequence 1, Appl
13	70.2	1.9	1392	4	US-09-412-554A-1	Sequence 1, Appl	Sequence 3, Appl
14	66.8	1.8	2949	4	US-09-412-554A-3	Sequence 3, Appl	Sequence 14, Appl
15	53.2	1.4	7218	1	US-08-232-463-14	Sequence 14, Appl	Sequence 14, Appl
16	50	1.4	7218	1	US-08-232-463-14	Sequence 14, Appl	Sequence 158, Appl
17	47	1.3	1771	4	US-09-907-794A-158	Sequence 158, Appl	Sequence 36, Appl
18	47	1.3	1771	4	US-09-866-028-36	Sequence 36, Appl	Sequence 158, Appl
19	47	1.3	1771	4	US-09-305-128A-158	Sequence 158, Appl	Sequence 86, Appl
20	47	1.3	1771	4	US-09-902-775A-158	Sequence 158, Appl	Sequence 2, Appl
21	45.8	1.2	11050	4	US-10-204-708-86	Sequence 86, Appl	Sequence 2, Appl
22	45.6	1.2	5852	1	US-07-667-106-2	Sequence 2, Appl	Sequence 1, Appl
23	45	1.2	6243	2	US-09-056-075-1	Sequence 1, Appl	Sequence 62, Appl
24	44.2	1.2	6801	4	US-10-204-708-62	Sequence 62, Appl	Sequence 22, Appl
25	44.2	1.2	11049	4	US-10-204-708-22	Sequence 22, Appl	Sequence 3, Appl
26	43.8	1.2	1813	3	US-09-071-224-3	Sequence 3, Appl	Sequence 47, Appl
27	43.8	1.2	5152	4	US-10-204-708-47	Sequence 47, Appl	

2521 TTTTCAGATTCAACAGGGGGAGAAATGGAGATTAAATCCTAAATCTGTACCTTGGGGAG 2580
2521 TTTTCAGATTCAACAGGGGGAGAAATGGAGATTAAATCCTAAATCTGTACCTTGGGGAG 2580
2581 GTCACTCAATTTACAGTTAGTCTGTGTCCTTTCAGACTCTGTGATTAATTAACCCCACTCAC 2640
2581 GTCACTCAATTTACAGTTAGTCTGTGTCCTTTCAGACTCTGTGATTAATTAACCCCACTCAC 2640
2641 TACCTCTGTTTCAGATGATTTGGAATACCAAGATTAAATCCTTTGACATAAGATCTCAAT 2700
2641 TACCTCTGTTTCAGATGATTTGGAATACCAAGATTAAATCCTTTGACATAAGATCTCAAT 2700
2701 TGCAGAAAGCAGATTAAAGACCATTCAAGAGAAATTAATTTAGTTGTAATGCAAGGCAA 2760
2701 TGCAGAAAGCAGATTAAAGACCATTCAAGAGAAATTAATTTAGTTGTAATGCAAGGCAA 2760
2761 CTGTGAGAAATCTGTGTGGCCAAAATAGAAATCCTCTAGTTTCTTCTGTTCTCAATTTGA 2820
2761 CTGTGAGAAATCTGTGTGGCCAAAATAGAAATCCTCTAGTTTCTTCTGTTCTCAATTTGA 2820
2821 AAGGAGAAATTCACATTTGTTTGTAGCAATTTCAAGCTTTTATGATCCATCCCATCTAAA 2880
2821 AAGGAGAAATTCACATTTGTTTGTAGCAATTTCAAGCTTTTATGATCCATCCCATCTAAA 2880
2881 ACTCTTCAAACTCCACTTGTTCAGTCTGAAATGAGCTCCCTGTCGAAGTCCCTTGGAGA 2940
2881 ACTCTTCAAACTCCACTTGTTCAGTCTGAAATGAGCTCCCTGTCGAAGTCCCTTGGAGA 2940
2941 ACTCACAGCAGCAGCCCTTAATCAAGAGTTTACCAGCCCTTGGACACTATGGAGGAGG 3000
2941 ACTCACAGCAGCAGCCCTTAATCAAGAGTTTACCAGCCCTTGGACACTATGGAGGAGG 3000
3001 GCAGAGTACACCAATTTGTTTAAAGCAAGAAACCAAGTGTCTTCACTAGTCAATTTA 3060
3001 GCAAGAGTACACCAATTTGTTTAAAGCAAGAAACCAAGTGTCTTCACTAGTCAATTTA 3060
3061 GAACATGGTTATCATCCAAGACTACTCTACCTGCAACATTGAACTCCCAAGAGCAAAATC 3120
3061 GAACATGGTTATCATCCAAGACTACTCTACCTGCAACATTGAACTCCCAAGAGCAAAATC 3120
3121 CACATTCCTCTGTAGTTCTGCAGCTTCTGTGTAATAGGCGAGCTGTCGCTATGCCGTA 3180
3121 CACATTCCTCTGTAGTTCTGCAGCTTCTGTGTAATAGGCGAGCTGTCGCTATGCCGTA 3180
3181 GAATCATGATCTGAGGACCAATTCATGGAAGCTGCTAAATAGCTAGTCTGGGGAGTCT 3240
3181 GAATCATGATCTGAGGACCAATTCATGGAAGCTGCTAAATAGCTAGTCTGGGGAGTCT 3240
3241 TCCATAAAGTTTTCGATGGAGCAAAACCAAGGATTAACTAGGTTTGGTTCCTTCAGCC 3300
3241 TCCATAAAGTTTTCGATGGAGCAAAACCAAGGATTAACTAGGTTTGGTTCCTTCAGCC 3300
3301 CTCTAAAGCATAGGGCTTAGCTTGCAGGCTTCTTGGGCTTCTCTGTGTGTAGTTT 3360
3301 CTCTAAAGCATAGGGCTTAGCTTGCAGGCTTCTTGGGCTTCTCTGTGTGTAGTTT 3360
3361 TGTAAACACTATAGCAATCTGTATAGATCCAGTGTCCATGMAACCTTCCCAATGCGGTG 3420
3361 TGTAAACACTATAGCAATCTGTATAGATCCAGTGTCCATGMAACCTTCCCAATGCGGTG 3420
3421 ACTCTGGAATATACAGTTTGGAAAGCAGGGTCTCTGCTGCTGCTAAACAGGCCACGT 3480
3421 ACTCTGGAATATACAGTTTGGAAAGCAGGGTCTCTGCTGCTGCTAAACAGGCCACGT 3480
3481 GGACCACTGAATGTCTTCCCTTACACCTATGTTTAAAGTAGTCAAACTTCAAGAAA 3540
3481 GGACCACTGAATGTCTTCCCTTACACCTATGTTTAAAGTAGTCAAACTTCAAGAAA 3540
3541 CAATCTAACCAAGTTCTGTGTCATATGTTTGTGAACTTGTATTTAGTAGG 3600
3541 CAATCTAACCAAGTTCTGTGTCATATGTTTGTGAACTTGTATTTAGTAGG 3600
3601 TTTCTATATTGCAATTAACCTGTTTGTGTAACCTGATTTCTTCTTTTCGATACTATTG 3660

Db 3601 TTTCTATATTGCAATTAACCTGTTTGTGTAACCTGATTTCTTTTCGATACTATTG 3660
QY 3661 ATGAATAAGAAATTAAGTGAATAAAAAA 3693
Db 3661 ATGAATAAGAAATTAAGTGAATAAAAAA 3693
RESULT 2
US-09-621-976-15155
; Sequence 15155, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 15155
; LENGTH: 896
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-15155
Query Match 24.2%; Score 892.8; DB 4; Length 896;
Best Local Similarity 99.8%; Pred. No. 2.1e-243;
Matches 894; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2668 CCAAGATTAAATCCTTGACATAGATCTCATTTGCAGAAAGCAGATTAAAGACCATCAG 2727
Db 1 CCAAGATTAAAGCCTTGACATAGATCTCATTTGCAGAAAGCAGATTAAAGACCATCAG 60
QY 2728 AAGGAAATTAATTTAGGTTGTAATGCAGGCAACTGTGAGAAACTGTTGTGCCAAAATA 2787
Db 61 AAGGAAATTAATTTAGGTTGTAATGCAGGCAACTGTGAGAAACTGTTGTGCCAAAATA 120
QY 2788 GAATTCCTTCTAGTTTCTTGTCTCATTTGAAAGGAGAAATTCCACTTGTTTAGCA 2847
Db 121 GAATTCCTTCTAGTTTCTTGTCTCATTTGAAAGGAGAAATTCCACTTGTTTAGCA 180
QY 2848 TTTCAAGCTTTTATGATCCATCCCATCTAAAACTCTTCAAACTCCACTTGTTCAGTCT 2907
Db 181 TTTCAAGCTTTTATGATCCATCCCATCTAAAACTCTTCAAACTCCACTTGTTCAGTCT 240
QY 2908 GAATTCGAGCTCCCTGTCGAGTCCCTTGGAGAACTCAGCAGCAGCCCTTAATCAAAG 2967
Db 241 GAATTCGAGCTCCCTGTCGAGTCCCTTGGAGAACTCAGCAGCAGCCCTTAATCAAAG 300
QY 2968 GTTTTACCAGCCCTTGGACACTATGAGGAGGAGGCAAGAGTACACCAATTTGTTAAAGC 3027
Db 301 GTTTTACCAGCCCTTGGACACTATGAGGAGGAGGCAAGAGTACACCAATTTGTTAAAGC 360
QY 3028 AAGAAACCAAGTGTCTCTTCACTAGTCAATTTAGACATGTTATCATCCAGACTATC 3087
Db 361 AAGAAACCAAGTGTCTCTTCACTAGTCAATTTAGACATGTTATCATCCAGACTATC 420
QY 3088 TACCCTGCAACATTCGAATCTCCCAAGGCAAAATCCCACTTCTTGAAGTCTTGCAAGCTTC 3147
Db 421 TACCCTGCAACATTCGAATCTCCCAAGGCAAAATCCCACTTCTTGAAGTCTTGCAAGCTTC 480
QY 3148 TGTGTAATAGGCGAGCTGTGCTGTATGCGGTAGAAATCAATGATCTGAGGACCATTCAT 3207
Db 481 TGTGTAATAGGCGAGCTGTGCTGTATGCGGTAGAAATCAATGATCTGAGGACCATTCAT 540
QY 3208 GGAAGCTGCTAAATAGCCTAGTCTCGGGAGTCTTCCATAAAGTCTTGTGATGGAGCAACA 3267
Db 541 GGAAGCTGCTAAATAGCCTAGTCTCGGGAGTCTTCCATAAAGTCTTGTGATGGAGCAACA 600
QY 3268 AACAGGATTAAACTAGGTTTGGTTCCTTCAGCCCTCTCTAAAGCATAGGGCTTAGCCCTGCA 3327

601 AACAGATTAACTAGGTTGGTTCTTCAGCCCTCTAAAGCATAGGGCTTAGCCTGA 660
 3328 GGCTTCCTGGGCTTCTCTGTGTGTAGTTTGTAAACACTATAGCATCTGTTAAGAT 3387
 661 GGCTTCCTGGGCTTCTCTGTGTGTAGTTTGTAAACACTATAGCATCTGTTAAGAT 720
 3388 CCAGTGTCCATGGAACCTTCCACATGCGTGCACATCTCGACATATATCAATGTTTGGAAA 3447
 721 CCAGTGTCCATGGAACCTTCCACATGCGTGCACATCTCGACATATATCAATGTTTGGAAA 780
 3448 GCAGGGTTCCTCTGCGCTCTAAACAGCCACGCGGACCACTGTGAATGCTTCTCTTAC 3507
 781 GCAGGGTTCCTCTGCGCTCTAAACAGCCACGCGGACCACTGTGAATGCTTCTCTTAC 840
 3508 ACCTATGTTTTTAAAGTAGTCAAACTTCAAGAAACAATCAAGTCTTCTGTTC 3563
 841 ACCTATGTTTTTAAAGTAGTCAAACTTCAAGAAACAATCAAGTCTTCTGTTC 896

RESULT 3

JS-09-643-597-209
 ; Sequence 209, Application US/09643597
 ; Patent No. 6426072
 ; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong
 ; APPLICANT: Fan, Liqun
 ; APPLICANT: Kalos, Michael D.
 ; APPLICANT: Boskur, Chaitanya S.
 ; APPLICANT: Hosken, Nancy
 ; APPLICANT: Fanger, Gary R.
 ; APPLICANT: Li, Samuel X.
 ; APPLICANT: Wang, Aijun
 ; APPLICANT: Skeiky, Iasir A.W.
 ; APPLICANT: Henderson, Robert A.
 ; APPLICANT: McNeill, Patricia D.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
 ; FILE REFERENCE: 210121.455C1
 ; CURRENT APPLICATION NUMBER: US/09/643,597
 ; CURRENT FILING DATE: 2000-08-21
 ; NUMBER OF SEQ ID NOS: 369
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 209
 ; LENGTH: 621
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-643-597-209

Query Match 16.8%; Score 619.4; DB 4; Length 621;
 Best Local Similarity 99.8%; Pred. No. 8.4e-166;
 Matches 620; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 3055 CATTAGAACATGGTTATCATCCAGACTACTCTACCCCTGCAACATTGAACTCCCAAGAG 3114
 1 CATTAGAACATGGTTATCATCCAGACTACTCTACCCCTGCAACATTGAACTCCCAAGAG 60
 3115 CAAATCCACATTCCTCTTGAGTTCTGCAAGTTCTGTGTAATAGGGCAGCTGCTCTAT 3174
 61 CAAATCCACATTCCTCTTGAGTTCTGCAAGTTCTGTGTAATAGGGCAGCTGCTCTAT 120
 3175 CCGTAGAATCACATGATCTGAGACCATTCATGGAAGCTCTAAATAGAGCTAGTCTGG 3234
 121 CCGTAGAATCACATGATCTGAGACCATTCATGGAAGCTCTAAATAGAGCTAGTCTGG 180
 3235 GAGTCTTCCATAAAGTTTGTGATGGAGCAACAAACAGGATTAACCTAGGTTGGTTCCT 3294
 181 GAGTCTTCCATAAAGTTTGTGATGGAGCAACAAACAGGATTAACCTAGGTTGGTTCCT 240
 3295 TCAGCCCTCTAAAGCATAGGGCTTAGGCTGAGGCTTCTTGGGCTTCTCTGTGTG 3354
 241 TCAGCCCTCTAAAGCATAGGGCTTAGGCTGAGGCTTCTTGGGCTTCTCTGTGTG 300

3355 TAGTTTGTAAACACTATAGCATCTGTTAAGATCCAGTGTCCATGGAACCTTCCACAT 3414
 301 TAGTTTGTAAACACTATAGCATCTGTTAAGATCCAGTGTCCATGGAACCTTCCACAT 360
 3415 GCGTGTACTGTGACTATATCACTGTTTGGAAAGCAGGGTTCCTCTGCGCTGCTAAACAGC 3474
 361 GCGTGTACTGTGACTATATCACTGTTTGGAAAGCAGGGTTCCTCTGCGCTGCTAAACAGC 420
 3475 CCAGTGTGACAGCTCTGAATGTCTTCTCTTTACACCTATGTTTAAAGTAGTCAAACTTC 3534
 421 CCAGTGTGACAGCTCTGAATGTCTTCTCTTTACACCTATGTTTAAAGTAGTCAAACTTC 480
 3535 AAGAAACAATCTAAACAAGTTCTGTTGATATGTTTGTGAACCTTGTATTGTATTATTA 3594
 481 AAGAAACAATCTAAACAAGTTCTGTTGATATGTTTGTGAACCTTGTATTGTATTATTA 540
 3595 GTAGGCTTCTATATTCATTTAACTTCTTCTTTTGTAACTCTGATTCCTTTTCGGATA 3654
 541 GTAGGCTTCTATATTCATTTAACTTCTTCTTTTGTAACTCTGATTCCTTTTCGGATA 600
 3655 CTATTGATGAATAAAGAAAT 3675
 601 CTATTGATGAATAAAGAAAT 621

RESULT 4

US-09-480-884A-209
 ; Sequence 209, Application US/09480884A
 ; Patent No. 6482597
 ; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong
 ; APPLICANT: Fan, Liqun
 ; APPLICANT: Kalos, Michael D.
 ; APPLICANT: Hosken, Nancy A.
 ; APPLICANT: Fanger, Gary R.
 ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
 ; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
 ; FILE REFERENCE: 210121.455C6
 ; CURRENT APPLICATION NUMBER: US/09/480,884A
 ; CURRENT FILING DATE: 2001-08-27
 ; NUMBER OF SEQ ID NOS: 330
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 209
 ; LENGTH: 621
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-480-884A-209

Query Match 16.8%; Score 619.4; DB 4; Length 621;
 Best Local Similarity 99.8%; Pred. No. 8.4e-166;
 Matches 620; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 3055 CATTAGAACATGGTTATCATCCAGACTACTCTACCCCTGCAACATTGAACTCCCAAGAG 3114
 1 CATTAGAACATGGTTATCATCCAGACTACTCTACCCCTGCAACATTGAACTCCCAAGAG 60
 3115 CAAATCCACATTCCTCTTGAGTTCTGCAAGTTCTGTGTAATAGGGCAGCTGCTCTAT 3174
 61 CAAATCCACATTCCTCTTGAGTTCTGCAAGTTCTGTGTAATAGGGCAGCTGCTCTAT 120
 3175 CCGTAGAATCACATGATCTGAGGACCATTCATGGAAGCTCTAAATAGAGCTAGTCTGG 3234
 121 CCGTAGAATCACATGATCTGAGGACCATTCATGGAAGCTCTAAATAGAGCTAGTCTGG 180
 3235 GAGTCTTCCATAAAGTTTGTGATGGAGCAACAAACAGGATTAACCTAGGTTGGTTCCT 3294
 181 GAGTCTTCCATAAAGTTTGTGATGGAGCAACAAACAGGATTAACCTAGGTTGGTTCCT 240
 3295 TCAGCCCTCTAAAGCATAGGGCTTAGGCTGAGGCTTCTTGGGCTTCTCTGTGTG 3354
 241 TCAGCCCTCTAAAGCATAGGGCTTAGGCTGAGGCTTCTTGGGCTTCTCTGTGTG 300
 3355 TAGTTTGTAAACACTATAGCATCTGTTAAGATCCAGTGTCCATGGAACCTTCCACAT 3414

Db 301 TAGTTTGTAAACACTATAGCATCTGTTAAGATCCAGTGCTCCATGGAACCTTCCACAT 360
2Y 3415 GCCGTGACTCTGGACTATATCAGTTTTTGGAAAGCAGGGTTCCTCTGCTGCTTAACAAGC 3474
2b 361 GCCGTGACTCTGGACTATATCAGTTTTTGGAAAGCAGGGTTCCTCTGCTGCTTAACAAGC 420
2Y 3475 CCAGTGGACCAAGTCTGAATGCTTCTCTTTACACCTATGTTTTTAAGTAGTCAAACTTC 3534
Db 421 CCAGTGGACCAAGTCTGAATGCTTCTCTTTACACCTATGTTTTTAAGTAGTCAAACTTC 480
QY 3535 AAGAAACAATCTAAACAAGTTTCTGTGTCATATGTTTGTGAACCTTGTAATTTGTAATTA 3594
Db 481 AAGAAACAATCTAAACAAGTTTCTGTGTCATATGTTTGTGAACCTTGTAATTTGTAATTA 540
QY 3595 GTAGCTTCTATATGCAATTTAACTTGTTTGTGAACCTTGTAATTTGTAATTTGTAATTA 3654
Db 541 GTAGCTTCTATATGCAATTTAACTTGTTTGTGAACCTTGTAATTTGTAATTTGTAATTA 600
QY 3655 CTATTGATGAATAAAGAAATT 3675
Db 601 CTATTGATGAATAAAGAAATT 621

RESULT 5
US-09-542-615A-209
; Sequence 209, Application US/09542615A
; Patent No. 6518256
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Ligu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C8
; CURRENT APPLICATION NUMBER: US/09/542,615A
; CURRENT FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 209
; LENGTH: 621
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-542-615A-209

Query Match 16.8%; Score 619.4; DB 4; Length 621;
Best Local Similarity 99.8%; Pred. No. 8.4e-166;
Matches 620; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 3055 CATTAGAACATGGTTATCATCAAGACTACTCTACCTGCAACATTGAACCTCCCAAGAG 3114
Db 1 CATTAGAACATGGTTATCATCAAGACTACTCTACCTGCAACATTGAACCTCCCAAGAG 60
QY 3115 CAAATCCACATTCCTCTGAGTTCTGCAGCTTCTGTGTAATAGGCGAGCTGCTCTAT 3174
Db 61 CAAATCCACATTCCTCTGAGTTCTGCAGCTTCTGTGTAATAGGCGAGCTGCTCTAT 120
QY 3175 GCCGTAGATCAATGATCTGAGGACCATTCATGGAAGCTGCTAAATAGCCTAGTCTGG 3234
Db 121 GCCGTAGATCAATGATCTGAGGACCATTCATGGAAGCTGCTAAATAGCCTAGTCTGG 180
QY 3235 GAGTCTTCCATAAAGTTTGCATGGAGCAACAACAGGATTAACCTAGTTTGGTTCCT 3294
Db 181 GAGTCTTCCATAAAGTTTGCATGGAGCAACAACAGGATTAACCTAGTTTGGTTCCT 240
QY 3295 TCAGCCCTCTAAAGCATAGGGCTTAGCCTGCAAGGCTTCCTTGGGCTTCTCTGTGTGTG 3354
Db 241 TCAGCCCTCTAAAGCATAGGGCTTAGCCTGCAAGGCTTCCTTGGGCTTCTCTGTGTGTG 300
QY 3355 TAGTTTGTAAACACTATAGCATCTGTTAAGATCCAGTGCTCCATGGAACCTTCCACAT 3414

Db 301 TAGTTTGTAAACACTATAGCATCTGTTAAGATCCAGTGCTCCATGGAACCTTCCACAT 360
QY 3415 GCCGTGACTCTGGACTATATCAGTTTTTGGAAAGCAGGGTTCCTCTGCTGCTTAACAAGC 3474
Db 361 GCCGTGACTCTGGACTATATCAGTTTTTGGAAAGCAGGGTTCCTCTGCTGCTTAACAAGC 420
QY 3475 CCAGTGGACCAAGTCTGAATGCTTCTCTTTACACCTATGTTTTTAAGTAGTCAAACTTC 3534
Db 421 CCAGTGGACCAAGTCTGAATGCTTCTCTTTACACCTATGTTTTTAAGTAGTCAAACTTC 480
QY 3535 AAGAAACAATCTAAACAAGTTTCTGTGTCATATGTTTGTGAACCTTGTAATTTGTAATTA 3594
Db 481 AAGAAACAATCTAAACAAGTTTCTGTGTCATATGTTTGTGAACCTTGTAATTTGTAATTA 540
QY 3595 GTAGCTTCTATATGCAATTTAACTTGTTTGTGAACCTTGTAATTTGTAATTTGTAATTA 3654
Db 541 GTAGCTTCTATATGCAATTTAACTTGTTTGTGAACCTTGTAATTTGTAATTTGTAATTA 600
QY 3655 CTATTGATGAATAAAGAAATT 3675
Db 601 CTATTGATGAATAAAGAAATT 621

RESULT 6
US-09-606-421B-209
; Sequence 209, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Ligu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606,421B
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 209
; LENGTH: 621
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-606-421B-209

Query Match 16.8%; Score 619.4; DB 4; Length 621;
Best Local Similarity 99.8%; Pred. No. 8.4e-166;
Matches 620; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 3055 CATTAGAACATGGTTATCATCAAGACTACTCTACCTGCAACATTGAACCTCCCAAGAG 3114
Db 1 CATTAGAACATGGTTATCATCAAGACTACTCTACCTGCAACATTGAACCTCCCAAGAG 60
QY 3115 CAAATCCACATTCCTCTGAGTTCTGCAGCTTCTGTGTAATAGGCGAGCTGCTCTAT 3174
Db 61 CAAATCCACATTCCTCTGAGTTCTGCAGCTTCTGTGTAATAGGCGAGCTGCTCTAT 120
QY 3175 GCCGTAGATCAATGATCTGAGGACCATTCATGGAAGCTGCTAAATAGCCTAGTCTGG 3234
Db 121 GCCGTAGATCAATGATCTGAGGACCATTCATGGAAGCTGCTAAATAGCCTAGTCTGG 180
QY 3235 GAGTCTTCCATAAAGTTTGCATGGAGCAACAACAGGATTAACCTAGTTTGGTTCCT 3294
Db 181 GAGTCTTCCATAAAGTTTGCATGGAGCAACAACAGGATTAACCTAGTTTGGTTCCT 240
QY 3295 TCAGCCCTCTAAAGCATAGGGCTTAGCCTGCAAGGCTTCCTTGGGCTTCTCTGTGTGTG 3354

Db 241 TCAGCCCTCTAAAAGCATAGGGCTTAGCCTGCAGGCTTCCCTTGGGCTTTCTCTGTGTGTG 300
 QY 3355 TAGTTTGTGTAACACTATAGCATCTGTAAAGATCCAGTGTCCATGGAACCTTCCACAT 3414
 Db 301 TAGTTTGTGTAACACTATAGCATCTGTAAAGATCCAGTGTCCATGGAACCTTCCACAT 360
 QY 3415 GCGGTGACTCTGGACTATATCAAGTTTGTGGAAGCAGGGTTCTCTGTGCTGTCAACAGC 3474
 Db 361 GCGGTGACTCTGGACTATATCAAGTTTGTGGAAGCAGGGTTCTCTGTGCTGTCAACAGC 420
 QY 3475 CCAGGTGGACCACTGCTGAATGCTTCTTCCCTTACACTATGTTTAAAGTAGTCAACCTC 3534
 Db 421 CCAGGTGGACCACTGCTGAATGCTTCTTCCCTTACACTATGTTTAAAGTAGTCAACCTC 480
 QY 3535 AGAAACAATCTAAACAAGTTTCTGTGTGATGTTTGTGGAACCTTGTGTAATTTA 3594
 Db 481 AGAAACAATCTAAACAAGTTTCTGTGTGATGTTTGTGGAACCTTGTGTAATTTA 540
 QY 3595 GTAGGCTTCTATATTCATTAACCTTGTGTTTGTGTAACCTTCTTCCCTTTCGGATA 3654
 Db 541 GTAGGCTTCTATATTCATTAACCTTGTGTTTGTGTAACCTTCTTCCCTTTCGGATA 600
 QY 3655 CTATTGATGAATAAGAAATT 3675
 Db 601 CTATTGATGAATAAGAAATT 621

RESULT 7

US-08-896-164-34/c
 ; Sequence 34, Application US/08896164
 ; Patent No. 6218521

GENERAL INFORMATION:

APPLICANT: OBATA, Yuichi
 TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ASSOCIATED
 TITLE OF INVENTION: WITH GASTRIC CANCER AND METHODS FOR
 TITLE OF INVENTION: DIAGNOSING AND TREATING GASTRIC CANCER
 NUMBER OF SEQUENCES: 87
 CORRESPONDENCE ADDRESS:
 ADDRESS: Felfe & Lynch
 STREET: 805 Third Avenue
 CITY: New York City
 STATE: New York
 COUNTRY: USA
 ZIP: 10022

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb

COMPUTER:

IBM PS/2

OPERATING SYSTEM:

PC-DOS

SOFTWARE:

Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/896,164

FILING DATE: July 17, 1997

CLASSIFICATION:

424

ATTORNEY/AGENT INFORMATION:

NAME: No. 6218521man D. Hanson

REGISTRATION NUMBER: 30,946

REFERENCE/DOCKET NUMBER: LUD 5499 - JEL/NDH/SLH

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 688-9200

TELEFAX: (212) 838-3884

INFORMATION FOR SEQ ID NO:

34:

SEQUENCE CHARACTERISTICS:

LENGTH: 708 base pairs

TYPE: nucleic acid

STRANDEDNESS: Double

TOPOLOGY: linear

US-08-896-164-34

Query Match 12.9% Score 477.2; DB 3; Length 708;

Best Local Similarity 80.8%; Pred. No. 2.4e-125;

Matches 573; Conservative 0; Mismatches 131; Indels 5; Gaps 5;

QY 2974 CCAGCCCTTGACACTATGAGGAGGGCAAGAGTACACC-AATTGTTAAAGCAAGAA 3032

Db 706 CCAGCCCTTGAAAAAATAAGGGGGGGGCAAGAAACACCCNNATTTGTTAAAAAGCAAG 647
 QY 3033 ACCACAGTGTCTCTTCACTAGTCAATTTAGAAACATGTTATCATCCAAAGACTACTCTACCC 3092
 Db 646 AACACAGGTTTTTTCATAGTATGTTAGAAAAGGTTTNNATCCAAGA-TAATTNACCC 588
 QY 3093 TGCAACATTGAATCCCAAGAGCAAAATCCACATTCCTCTTGAGTTCTTGAGCTTCTGTGT 3152
 Db 587 GGNAAACATTG-ANTCCNAGNGCNAATTCACATTCCTGAGTTCGAGGTTCCGNGT 529
 QY 3153 AAATAGGGCAGTGTCTGTCTATCCCGTAGAATCACATGATCTGAGGACCATTCATCGAAG 3212
 Db 528 AAATAGGGCAGTGTCTGTCTATCCCGTAGAATCACATGATCTGAGGACCATTCATCGAAG 469
 QY 3213 CTGCTAAATAGCCCTAGTGTGGGAGTCTTCCATAAAGTTTTCATGAGGACCAACAACAG 3272
 Db 468 NNGNTAAATAGCCCTAGTGTGGG-GTTTTCATATAAAGTTTTCATGAGGACCAACAACAG 410
 QY 3273 GATTAACTAGGTTTGGTTCCTTCAGCCCTCTATAAAGCATAGGCTTAGCCTGAGGCTT 3332
 Db 409 GNTTAAATAGGTTTGGTTCCTTCAGCCCTCTATAAAGCATAGGCTTAGCCTGAGG-IT 351
 QY 3333 CCTTGGCTTTCTCTGTGTGTGTGTTTGTGTAACACTATAGCATCTGTTAAGATCCAGT 3392
 Db 350 CCTTGGGTTTTTTTNGGGGNGTAGTTTGTAAANANTATAGCATTTGTTAAGATNCAGN 291
 QY 3393 GTCATGGAACCTTCCCATGCTCCGCTGACTATATCAGTTTTCGAAAGCAGG 3452
 Db 290 GTCATGGAACCTTCCCATGCTCCGCTGACTATATCAGTTTTCGAAAGCAGG 231
 QY 3453 GTTCTCTGCTCTAAACAGCCAGCTGAGCAGTCTGAATGCTTCTTCTTTACACCTA 3512
 Db 230 GTTCTTTCCTGCTGTAANAAGCCCAAGTGCAGCAGTTTGNANGTTTTTCTTTANACCTA 171
 QY 3513 TGTTTTAAGTACTCAACTTCAAGAAACAATCTAAACAGTTTCTGTTGCATATGTT 3572
 Db 170 TGNTTTTAAGTAGTNNAAATTAAGGAACNAITTAACCAAGTTTTTGTGCATANGTT 111
 QY 3573 TGTGAATCTGTTATTTGTTATTTAGTAGCTTCTATATTGCAATTAACCTGTTTTTGAAC 3632
 Db 110 TGIGNAATTCGTTATTTGTTATTTAGTAGGTTTTTATATTGCANTTAATTTGTTTGTAAAT 51
 QY 3633 CTGATTTCTTCTTTCGGATCTATTTGATGAATAAGAAATTAAGTG 3681
 Db 50 CCTGATTTTCTTTCGGATCTATTTGATGAATAAGAAATTAAGTG 2

RESULT 8

US-09-643-597-315

; Sequence 315, Application US/09643597

; Patent No. 6426072

GENERAL INFORMATION:

APPLICANT: Wang, Tongtong

APPLICANT: Fan, Liqun

APPLICANT: Kalos, Michael D.

APPLICANT: Bangur, Chaitanya S.

APPLICANT: Hosken, Nancy

APPLICANT: Fanger, Gary R.

APPLICANT: Li, Samuel X.

APPLICANT: Wang, Aijun

APPLICANT: Skeiky, Yasir A.W.

APPLICANT: Henderson, Robert A.

APPLICANT: McNeill, Patricia D.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER

FILE REFERENCE: 210121.455C11

CURRENT APPLICATION NUMBER: US/09/643,597

CURRENT FILING DATE: 2000-08-21

NUMBER OF SEQ ID NOS: 369

SOFTWARE: Fast-Seq for Windows Version 3.0

SEQ ID NO 315

LENGTH: 336

TYPE: DNA
ORGANISM: Homo sapien
US-09-480-884A-315
Query Match 9.1%; Score 336; DB 4; Length 336;
Best Local Similarity 100.0%; Pred. No. 2e-85;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3057 TTTAGAACATGGTTATCATCCAAAGACTACTCTACCCCTGCAACATTGAATCCCAAGAGCA 3116
Db 1 TTTAGAACATGGTTATCATCCAAAGACTACTCTACCCCTGCAACATTGAATCCCAAGAGCA 60
QY 3117 AATCCACATTCCTCTTGAGTTCTGAGCTTCTGAGTAAATAGGCGAGCTGTCGTCTATGC 3176
Db 61 AATCCACATTCCTCTTGAGTTCTGAGTAAATAGGCGAGCTGTCGTCTATGC 120
QY 3177 CGTAGAATCACATGATCTGAGGACCATTCATGGAAGCTGCTAAATAGCTAGCTGGGA 3236
Db 121 CGTAGAATCACATGATCTGAGGACCATTCATGGAAGCTGCTAAATAGCTAGCTGGGA 180
QY 3237 GTCTTCCATAAAGTTTTCATGGAGCAACAAACAGGATTAAATAGCTTTGGTTCCCTTC 3296
Db 181 GTCTTCCATAAAGTTTTCATGGAGCAACAAACAGGATTAAATAGCTTTGGTTCCCTTC 240
QY 3297 AGCCCTCTAAAGCATAGGGCTTAGCCTGCGAGGCTTCTGGGCTTCTCTGTGTGTGTA 3356
Db 241 AGCCCTCTAAAGCATAGGGCTTAGCCTGCGAGGCTTCTGGGCTTCTCTGTGTGTGTA 300
QY 3357 GTTTTGTAAACACTATAGCATCTGTTAAGATCCAGT 3392
Db 301 GTTTTGTAAACACTATAGCATCTGTTAAGATCCAGT 336

RESULT 9
US-09-480-884A-315
; Sequence 315, Application US/09480884A
; Patent No. 6482597
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; FILE REFERENCE: 210121.455C6
; CURRENT APPLICATION NUMBER: US/09/480,884A
; NUMBER OF SEQ ID NOS: 330
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 315
; LENGTH: 336
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-480-884A-315
Query Match 9.1%; Score 336; DB 4; Length 336;
Best Local Similarity 100.0%; Pred. No. 2e-85;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3057 TTTAGAACATGGTTATCATCCAAAGACTACTCTACCCCTGCAACATTGAATCCCAAGAGCA 3116
Db 1 TTTAGAACATGGTTATCATCCAAAGACTACTCTACCCCTGCAACATTGAATCCCAAGAGCA 60
QY 3117 AATCCACATTCCTCTTGAGTTCTGAGCTTCTGTAATAGGCGAGCTGTCGTCTATGC 3176
Db 61 AATCCACATTCCTCTTGAGTTCTGAGCTTCTGTAATAGGCGAGCTGTCGTCTATGC 120
QY 3177 CGTAGAATCACATGATCTGAGGACCATTCATGGAAGCTGCTAAATAGCTAGCTGGGA 3236
Db 121 CGTAGAATCACATGATCTGAGGACCATTCATGGAAGCTGCTAAATAGCTAGCTGGGA 180
QY 3237 GTCTTCCATAAAGTTTTCATGGAGCAACAAACAGGATTAAATAGCTTTGGTTCCCTTC 3296

Db 181 GTCTTCCATAAAGTTTTCATGGAGCAACAAACAGGATTAAATAGCTTTGGTTCCCTTC 240
QY 3297 AGCCCTCTAAAGCATAGGGCTTAGCCTGCGAGGCTTCTGGGCTTCTCTGTGTGTGTA 3356
Db 241 AGCCCTCTAAAGCATAGGGCTTAGCCTGCGAGGCTTCTGGGCTTCTCTGTGTGTGTA 300
QY 3357 GTTTTGTAAACACTATAGCATCTGTTAAGATCCAGT 3392
Db 301 GTTTTGTAAACACTATAGCATCTGTTAAGATCCAGT 336

RESULT 10
US-09-542-615A-315
; Sequence 315, Application US/09542615A
; Patent No. 6518256
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; FILE REFERENCE: 210121.455C8
; CURRENT APPLICATION NUMBER: US/09/542,615A
; CURRENT FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 315
; LENGTH: 336
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-542-615A-315
Query Match 9.1%; Score 336; DB 4; Length 336;
Best Local Similarity 100.0%; Pred. No. 2e-85;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3057 TTTAGAACATGGTTATCATCCAAAGACTACTCTACCCCTGCAACATTGAATCCCAAGAGCA 3116
Db 1 TTTAGAACATGGTTATCATCCAAAGACTACTCTACCCCTGCAACATTGAATCCCAAGAGCA 60
QY 3117 AATCCACATTCCTCTTGAGTTCTGAGCTTCTGTGTAATAGGCGAGCTGTCGTCTATGC 3176
Db 61 AATCCACATTCCTCTTGAGTTCTGAGCTTCTGTGTAATAGGCGAGCTGTCGTCTATGC 120
QY 3177 CGTAGAATCACATGATCTGAGGACCATTCATGGAAGCTGCTAAATAGCTAGCTGGGA 3236
Db 121 CGTAGAATCACATGATCTGAGGACCATTCATGGAAGCTGCTAAATAGCTAGCTGGGA 180
QY 3237 GTCTTCCATAAAGTTTTCATGGAGCAACAAACAGGATTAAATAGCTTTGGTTCCCTTC 3296
Db 181 GTCTTCCATAAAGTTTTCATGGAGCAACAAACAGGATTAAATAGCTTTGGTTCCCTTC 240
QY 3297 AGCCCTCTAAAGCATAGGGCTTAGCCTGCGAGGCTTCTGGGCTTCTCTGTGTGTGTA 3356
Db 241 AGCCCTCTAAAGCATAGGGCTTAGCCTGCGAGGCTTCTGGGCTTCTCTGTGTGTGTA 300
QY 3357 GTTTTGTAAACACTATAGCATCTGTTAAGATCCAGT 3392
Db 301 GTTTTGTAAACACTATAGCATCTGTTAAGATCCAGT 336

RESULT 11
US-09-606-421B-315
; Sequence 315, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.

```

; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606.421B
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 315
; LENGTH: 336
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-606-421B-315

Query Match          9.1%; Score 336; DB 4; Length 336;
Best Local Similarity 100.0%; Pred. No. 2e-85;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3057 TTTAGAACATGGTTATATCATCCAGACTACTCTACCTGCAACATTGAATCCCAAGACCA 3116
DB 1 TTTAGAACATGGTTATATCATCCAGACTACTCTACCTGCAACATTGAATCCCAAGACCA 60

QY 3117 AATCCACATTCCTCTTGAGTCTTCAGCTTCTGTGTAATAGGCGAGCTGCGTCTATGC 3176
DB 61 AATCCACATTCCTCTTGAGTCTTCAGCTTCTGTGTAATAGGCGAGCTGCGTCTATGC 120

QY 3177 CGTAGAATCACATGATCTGAGGACCATTCATGGAAGCTGCTTAATAGCTAGTCTGGGGA 3236
DB 121 CGTAGAATCACATGATCTGAGGACCATTCATGGAAGCTGCTTAATAGCTAGTCTGGGGA 180

QY 3237 GCTTCCATTAAGTTTTCATGAGCAACAAACAGGATTAATAGTGTTCCTTC 3296
DB 181 GCTTCCATTAAGTTTTCATGAGCAACAAACAGGATTAATAGTGTTCCTTC 240

QY 3297 AGCCCTCTAAAAGCATAGGGCTTAGCCTGCAAGCTTCCTGGGCTTCTCTGTGTGTA 3356
DB 241 AGCCCTCTAAAAGCATAGGGCTTAGCCTGCAAGCTTCCTGGGCTTCTCTGTGTGTA 300

QY 3357 GTTTTGTAAACACTATAGCTCTGTTAGATCCAGT 3392
DB 301 GTTTTGTAAACACTATAGCTCTGTTAGATCCAGT 336

```

```

RESULT 12
US-09-412-554A-7
; Sequence 7, Application US/09412554A
; Patent No. 6355788
; GENERAL INFORMATION:
; APPLICANT: Conklin, Darrell
; APPLICANT: Ellsworth, Jeff L.
; TITLE OF INVENTION: FOLLISTATIN RELATED PROTEIN ZFSTA2
; FILE REFERENCE: 98-50
; CURRENT APPLICATION NUMBER: US/09/412,554A
; CURRENT FILING DATE: 1999-10-05
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 1255
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: 3' end of zfsta2 nucleotide sequence
US-09-412-554A-7

Query Match          1.9%; Score 70.2; DB 4; Length 1255;
Best Local Similarity 64.4%; Pred. No. 1.8e-09;
Matches 105; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

```

```

QY 214 CACAGAGAAAGGGGAACCCACCTGCTCTGATGAGCAATGCAAACTCACAAGAGGCC 273
DB 294 CAGAGAGACAGGGCAAGCAGAGATGCGCTGTATGACCTTTGCAAACTCCTACTACAAACC 353
QY 274 TGTGTGTGCGAGTATATGCAAGACCTACCTCAACCACTGTGAATCGAGATGCGCTG 333
DB 354 TGTGTGTGATCTGACGAGAAATCTATGAAAAACCACTGTGAAGTGCAAGAGCTGCTTG 413
QY 334 CTTCACTGGATCCAAAATCCAGGTTGATTAGATGGACACTGC 376
DB 414 CCTGAAAAAACAAGATGATACCAATGTTTCAATGAGACTGC 456

RESULT 13
US-09-412-554A-1
; Sequence 1, Application US/09412554A
; Patent No. 6355788
; GENERAL INFORMATION:
; APPLICANT: Conklin, Darrell
; APPLICANT: Ellsworth, Jeff L.
; TITLE OF INVENTION: FOLLISTATIN RELATED PROTEIN ZFSTA2
; FILE REFERENCE: 98-50
; CURRENT APPLICATION NUMBER: US/09/412,554A
; CURRENT FILING DATE: 1999-10-05
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 3192
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (58) ... (3006)
US-09-412-554A-1

Query Match          1.9%; Score 70.2; DB 4; Length 3192;
Best Local Similarity 64.4%; Pred. No. 3.3e-09;
Matches 105; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

```

```

QY 214 CACAGAGAAAGGGGAACCCACCTGCTCTGATGAGCAATGCAAACTCACAAGAGGCC 273
DB 294 CAGAGAGACAGGGCAAGCAGAGATGCGCTGTATGACCTTTGCAAACTCCTACTACAAACC 353
QY 274 TGTGTGTGCGAGTATATGCAAGACCTACCTCAACCACTGTGAATCGAGATGCGCTG 333
DB 354 TGTGTGTGATCTGACGAGAAATCTATGAAAAACCACTGTGAAGTGCAAGAGCTGCTTG 413
QY 334 CTTCACTGGATCCAAAATCCAGGTTGATTAGATGGACACTGC 376
DB 414 CCTGAAAAAACAAGATGATACCAATGTTTCAATGAGACTGC 456

```

```

RESULT 14
US-09-412-554A-3
; Sequence 3, Application US/09412554A
; Patent No. 6355788
; GENERAL INFORMATION:
; APPLICANT: Conklin, Darrell
; APPLICANT: Ellsworth, Jeff L.
; TITLE OF INVENTION: FOLLISTATIN RELATED PROTEIN ZFSTA2
; FILE REFERENCE: 98-50
; CURRENT APPLICATION NUMBER: US/09/412,554A
; CURRENT FILING DATE: 1999-10-05
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 2949
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Degenerate oligonucleotide sequence encoding the
; OTHER INFORMATION: zfsta2 polypeptide of SEQ ID NO:2.
; NAME/KEY: variation

```

LOCATION: (1)....(2949)
OTHER INFORMATION: Each N is independently any nucleotide.
US-09-412-554A-3

Query Match 1.8%; Score 66.8; DB 4; Length 2949;
Best Local Similarity 41.5%; Pred. No. 2.9e-08;
Matches 86; Conservative 42; Mismatches 76; Indels 3; Gaps 1;
2y 173 TGTCCCAATGTGTTTGTGGAGCGCGCGGAATGTG--CAGTCACAGAGAAGGGGA 229
Db 193 TGYGARAARATAYTGYGNYTNGMNGNCAYTGYTNAENMNGNGARANGNCAR 252
2y 230 CCCACCTGTCTCTCAATGAGCAATGCAACCTCACAAGAGGCGCTGTGTGTCAGTAAT 289
Db 253 GCNGAR:GYGCGTGYATGAYTNTGYAARMGNCAYTAYAACCGTNTGYGWNNGAY 312
Qy 290 GCGAAGACCTACCTACCACTGTGAAGTGCATGAGATGCGCTGCCTCACTGGATCCAAA 349
Db 313 GCGGARTTYTAYGARAAYCAYTGYGAGTNGAYNGCNGTGYTNAARAARCAPAAR 372
Qy 350 ATCCAGGTTGATTACGATGGACACTGC 376
Db 373 ATHACNATHGTCNAYAYGARGAATGY 399

RESULT 15
US-08-232-463-14/c
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232.463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZGPT-Fls
US-08-232-463-14

Query Match 1.4%; Score 53.2; DB 1; Length 7218;
Best Local Similarity 4.5%; Pred. No. 0.00039;
Matches 13; Conservative 172; Mismatches 105; Indels 0; Gaps 0;
Qy 770 GAGAAGAAGTGTGCGCTGGAGATGAACAGTATGCAGATGGAGCTGACACCGAGTGCAC 829
Db 1331 RRR 1272
Qy 830 TGTAAACCGCTGTGTCTGTGCTGTGAAATTTGGTCTGTACAGCCATGACCTGTGACGGA 889
Db 1271 RRR 1212
Qy 890 AAGAATCAAGAAGGGGCGCCAGACCCAGACAGAGAGAGATGACACGATATGTCCAGGAG 949
Db 1211 RRR 1152
Qy 950 CTCCTCAAGCATCAGAAACAGCTGAAAGACCAAGAGAGTGAAGCACCAGAGATCTTAA 1009
Db 1151 RRR 1092
Qy 1010 TGAGGAGGCACAGACAGTGTGTGATCCAGCATCTTCTTCCACTTTCAGC 1059
Db 1091 RRR 1042

Search completed: March 11, 2004, 14:41:33
Job time : 172.644 secs

181 TGTGTTTGTGAGCGCGCGGAATGTGCACTCAGAGAAAGGGAAACCACTGCT 240
 181 TGTGTTTGTGAGCGCGCGGAATGTGCACTCAGAGAAAGGGAAACCACTGCT 240
 241 CTGCAATGAGCAATCAAACTTCAAGAGGCGCTGTGTGCACTAATGCAAGACCTA 300
 241 CTGCAATGAGCAATCAAACTTCAAGAGGCGCTGTGTGCACTAATGCAAGACCTA 300
 301 CTTCAACCACTGTGAATGATCGAGATGCTGCTCACTGATCCAAATCCAGTTGA 360
 301 CTTCAACCACTGTGAATGATCGAGATGCTGCTCACTGATCCAAATCCAGTTGA 360
 361 TTACGATGACACTGCAAGAGAGAAATCCGTAGTCCATCTGACGCCAGTTGTTG 420
 361 TTACGATGACACTGCAAGAGAGAAATCCGTAGTCCATCTGACGCCAGTTGTTG 420
 421 CTATCAGTCCAACTGTGATGCTGCGAGTCCGATCATCCAGTGGCTGGAAGCTGAGAT 480
 421 CTATCAGTCCAACTGTGATGCTGCGAGTCCGATCATCCAGTGGCTGGAAGCTGAGAT 480
 481 CATCTCAGATGGCTGTTCTTAAGGCGAGCAACTACAGTGAATCCTAGACAAGTATT 540
 481 CATCTCAGATGGCTGTTCTTAAGGCGAGCAACTACAGTGAATCCTAGACAAGTATT 540
 541 TAAGAACTTTGATAATGGTGATTTCTGCGCTGGACTCCAGTGAATTCCTGAAGTTGTGA 600
 541 TAAGAACTTTGATAATGGTGATTTCTGCGCTGGACTCCAGTGAATTCCTGAAGTTGTGA 600
 601 ACAGATGAACCTGCATCAATTAACAGTATCCAGACCGAGGAGCAACAACTGCT 660
 601 ACAGATGAACCTGCATCAATTAACAGTATCCAGACCGAGGAGCAACAACTGCT 660
 661 TAGGGACTCTGTGTTGATGCTCTCAATGAACCTGTGATGAATATGCTGATGAAACT 720
 661 TAGGGACTCTGTGTTGATGCTCTCAATGAACCTGTGATGAATATGCTGATGAAACT 720
 721 CAGCTTCAAGAGTTCTCAAGTGCCTCAACCCATCTTTCAACCCCTCCTGAGAAGAGTG 780
 721 CAGCTTCAAGAGTTCTCAAGTGCCTCAACCCATCTTTCAACCCCTCCTGAGAAGAGTG 780
 781 TGCCCTGAGGATGAACGATATCAGATGGAGTGAAGCGAGTGGAGTCTGTAACCGCTG 840
 781 TGCCCTGAGGATGAACGATATCAGATGGAGTGAAGCGAGTGGAGTCTGTAACCGCTG 840
 841 TGTCTGCTGCTGGAATTTGGTCTGTACAGCCATGACCTGTGACGGAAGAAATCAGAA 900
 841 TGTCTGCTGCTGGAATTTGGTCTGTACAGCCATGACCTGTGACGGAAGAAATCAGAA 900
 901 GGGGCCCCAGCCAGACAGAGGAGGATGACAGATATGTCAGAGCTCCAAAGCA 960
 901 GGGGCCCCAGCCAGACAGAGGAGGATGACAGATATGTCAGAGCTCCAAAGCA 960
 961 TCAGGAAACAGCTGAAAGAACCAAGAGAGTGAGCAACCAAGAGATCTAATGAGGAGGCAC 1020
 961 TCAGGAAACAGCTGAAAGAACCAAGAGAGTGAGCAACCAAGAGATCTAATGAGGAGGCAC 1020
 1021 AGACCACTGTGATCCAGCATCTTCTCCATCTCAGCGCTGAGTTCACTACACAG 1080
 1021 AGACCACTGTGATCCAGCATCTTCTCCATCTCAGCGCTGAGTTCACTACACAG 1080
 1081 TGTCTGCTACAGTCCCAATCACCAGTATTTGCTTATATAGCAATGAGTTTATTTGT 1140
 1081 TGTCTGCTACAGTCCCAATCACCAGTATTTGCTTATATAGCAATGAGTTTATTTGT 1140
 1141 TTATTTGTTTGCATTAAGGATATGAGTGGCTGCTAGGAGGAGGAGGCGCACAGCC 1200
 1141 TTATTTGTTTGCATTAAGGATATGAGTGGCTGCTAGGAGGAGGAGGCGCACAGCC 1200
 1201 TTCAATTTCTAGAGTGTCTTAAGAGAACTGTAAATGGTCTCTGCGGCTGAGGCTAGT 1260
 1201 TTCAATTTCTAGAGTGTCTTAAGAGAACTGTAAATGGTCTCTGCGGCTGAGGCTAGT 1260

1261 AAGAACTGCATCAGATTGAAGAGGAAACAGACCAATCTGACCTCTTTTGAGTTT 1320
 1261 AAGAACTGCATCAGATTGAAGAGGAAACAGACCAATCTGACCTCTTTTGAGTTT 1320
 1321 ACTGCATCTGTACAGAGGCTGCAAGAGGAGTGCACGATCCAGAGAGAACTTAGCAGG 1380
 1321 ACTGCATCTGTACAGAGGCTGCAAGAGGAGTGCACGATCCAGAGAGAACTTAGCAGG 1380
 1381 GTCCCGGAGGAGAGGTTTGGGAGTCCACGAGGAGGAAAGCTCTCTGTTCCAGCTTC 1440
 1381 GTCCCGGAGGAGAGGTTTGGGAGTCCACGAGGAGGAAAGCTCTCTGTTCCAGCTTC 1440
 1441 TTTCCATTGCGGTGAGCATGACAGACTCCAGCATCCAGCATCTCTTGGTCCCAATAAC 1500
 1441 TTTCCATTGCGGTGAGCATGACAGACTCCAGCATCCAGCATCTCTTGGTCCCAATAAC 1500
 1501 TGCTCTAGATACATAGCCATCTGTAGTTAACTACAGGAGTCCCTCAGACTTCGATGGAG 1560
 1501 TGCTCTAGATACATAGCCATCTGTAGTTAACTACAGGAGTCCCTCAGACTTCGATGGAG 1560
 1561 TTTCTGGAGGAGTACACCCAAATGATCGAGTACTTGTATACCTTGGAGCCCTTAGCGAC 1620
 1561 TTTCTGGAGGAGTACACCCAAATGATCGAGTACTTGTATACCTTGGAGCCCTTAGCGAC 1620
 1621 CTAAACCAATTTTAAATACTTTTACCAAGAGTGTCTCTCTGTAACCACTTTT 1680
 1621 CTAAACCAATTTTAAATACTTTTACCAAGAGTGTCTCTCTGTAACCACTTTT 1680
 1681 TTTGGCAAGTGTACTTATTTCAATTAATTAATTAATTAATTAATTAATTAATTT 1740
 1681 TTTGGCAAGTGTACTTATTTCAATTAATTAATTAATTAATTAATTAATTAATTT 1740
 1741 ATTTCTGTAGTATAGCTTTTGTAAATTAATTTTCTAGTAGTCCCACTTTCATA 1800
 1741 ATTTCTGTAGTATAGCTTTTGTAAATTAATTTTCTAGTAGTCCCACTTTCATA 1800
 1801 GGTGGAAGGAGTTTGGGTTCTTCTGTCAGGGGCTGAATAAACCCAGATGTCCTCAC 1860
 1801 GGTGGAAGGAGTTTGGGTTCTTCTGTCAGGGGCTGAATAAACCCAGATGTCCTCAC 1860
 1861 CCTGCCACATACATAGATGCGAGCCATAGTTGGGCCCTTACGAGTCCAGCTATC 1920
 1861 CCTGCCACATACATAGATGCGAGCCATAGTTGGGCCCTTACGAGTCCAGCTATC 1920
 1921 TGCCAGAGGAGGAGGCTTACAGCTTAAGCCAGGGGAGGAGCATCTTCAATAAAAA 1980
 1921 TGCCAGAGGAGGAGGCTTACAGCTTAAGCCAGGGGAGGAGCATCTTCAATAAAAA 1980
 1981 CTTTCAAGATCCAAACATTAATTTTATTTTATTTTCTGAGAGTTGAGGCAATCAGT 2040
 1981 CTTTCAAGATCCAAACATTAATTTTATTTTATTTTCTGAGAGTTGAGGCAATCAGT 2040
 2041 ATTTCCAGGATGGGACAGGCGAGGCGAGGCTTAGGATATCCAGCTTACCAAT 2100
 2041 ATTTCCAGGATGGGACAGGCGAGGCGAGGCTTAGGATATCCAGCTTACCAAT 2100
 2101 ATGCTCATTTCACTAATCTAGGAGGAGTGGCCCTGTCTCTTTTCTTGGAGCTC 2160
 2101 ATGCTCATTTCACTAATCTAGGAGGAGTGGCCCTGTCTCTTTTCTTGGAGCTC 2160
 2161 AGTTTCTCAGTGTGTTGAAGATGCACTAATCTTTTGTATTTGATTAAGTTAATTC 2220
 2161 AGTTTCTCAGTGTGTTGAAGATGCACTAATCTTTTGTATTTGATTAAGTTAATTC 2220
 2221 TGTGTTCTGATCATTTGGTCCAGAGGAGATAGTTTCTGTGATTTTCTTCTTCTCT 2280
 2221 TGTGTTCTGATCATTTGGTCCAGAGGAGATAGTTTCTGTGATTTTCTTCTTCTCT 2280
 2281 ATAGATTAATTAATCTTGTACTAGAACAGAGAAATGTCAGATGGCCAAAAACAAGATG 2340
 2281 ATAGATTAATTAATCTTGTACTAGAACAGAGAAATGTCAGATGGCCAAAAACAAGATG 2340
 2341 ACCAGATTGATCTCAGGCTGATGACCTTACAGGCTGCTGATGATGATGAGTCTCTCATG 2400

ib	2341	ACCAGATTGATCTCAGCGCTGATGACCCCTACAGGTGCTGCTATGATATGGAATGCTCATG	2400
iy	2401	GGTAAGCAGGAAGAGATGGGAAAGAGAAACCAACCCACTCTGTCTTCATATTTTCATTT	2460
ib	2401	GGTAAGCAGGAAGAGATGGGAAAGAGAAACCAACCCACTCTGTCTTCATATTTTCATTT	2460
iy	2461	CATGTTTAACTCCCGCTGGAAATAGAAAGCATTCCTTTAGAGATGAGGATAAAAGAAAG	2520
ib	2461	CATGTTTAACTCCCGCTGGAAATAGAAAGCATTCCTTTAGAGATGAGGATAAAAGAAAG	2520
iy	2521	TTTTCAGATTCACACAGGGGGAAGAAATGAGATTTAATCCTTAAACTGTGACTTGGGAG	2580
ib	2521	TTTTCAGATTCACACAGGGGGAAGAAATGAGATTTAATCCTTAAACTGTGACTTGGGAG	2580
iy	2581	GTCAAGTCATTTCAGATTAGTCTGTGCTTTTCGACCTTCTGTGATTTATTAACCCCACTCAC	2640
ib	2581	GTCAAGTCATTTCAGATTAGTCTGTGCTTTTCGACCTTCTGTGATTTATTAACCCCACTCAC	2640
iy	2641	TACCCTGTTTTCAGATTCATTTTGAATACCAAGATTAATACTCTTGACATAAGATCTCATTT	2700
ib	2641	TACCCTGTTTTCAGATTCATTTTGAATACCAAGATTAATACTCTTGACATAAGATCTCATTT	2700
iy	2701	TGCAGAAACAGATTTAAAGACCATCAGAGAGAAATTTATTTAGTGTGTAATGACACGGCAA	2760
ib	2701	TGCAGAAACAGATTTAAAGACCATCAGAGAGAAATTTATTTAGTGTGTAATGACACGGCAA	2760
iy	2761	CTGTGAGAAACTGTCTGTGCCAAAAATAGAAATTCCTTCTAGTTTTCTGTCTTCATTTGA	2820
ib	2761	CTGTGAGAAACTGTCTGTGCCAAAAATAGAAATTCCTTCTAGTTTTCTGTCTTCATTTGA	2820
iy	2821	AAGGAGAAAATTCACATTTGTGTTTAGCATTTTCAAGCTTTTATGATTCATCCCATTAATA	2880
ib	2821	AAGGAGAAAATTCACATTTGTGTTTAGCATTTTCAAGCTTTTATGATTCATCCCATTAATA	2880
iy	2881	ACTCTTTCAAACTCCACTGTTTCAGTCTGAAATGCAGCTCCCTCTCCAAGTGCCTTCGAGA	2940
ib	2881	ACTCTTTCAAACTCCACTGTTTCAGTCTGAAATGCAGCTCCCTCTCCAAGTGCCTTCGAGA	2940
iy	2941	ACTCACAGCAGCAGCGCTTAATCAAAAGTTTTTACAGCCCTTGGACACTATGGGAGGAGG	3000
ib	2941	ACTCACAGCAGCAGCGCTTAATCAAAAGTTTTTACAGCCCTTGGACACTATGGGAGGAGG	3000
iy	3001	GCAAGAGTACACCAATTTGTTTAAAGCAAGAAACCAAGTGTCTCTTCACTAGTCATTTA	3060
ib	3001	GCAAGAGTACACCAATTTGTTTAAAGCAAGAAACCAAGTGTCTCTTCACTAGTCATTTA	3060
iy	3061	GAAACATGGTTATCATCCAGACTACTCTACCCCTGCACATATGAACCTCCCAAGACGAAATC	3120
ib	3061	GAAACATGGTTATCATCCAGACTACTCTACCCCTGCACATATGAACCTCCCAAGACGAAATC	3120
iy	3121	CACATTCCTCTTCAGTTCTCAGCTTCTGTGTAAATAGGGCAGCTGTCGTCTATGCCGTA	3180
ib	3121	CACATTCCTCTTCAGTTCTCAGCTTCTGTGTAAATAGGGCAGCTGTCGTCTATGCCGTA	3180
iy	3181	GAATCACATGATCTGAGGACATTCATFGGAAGCTGTCTAAATAGCCTAGTCTGGGAGTCT	3240
ib	3181	GAATCACATGATCTGAGGACATTCATFGGAAGCTGTCTAAATAGCCTAGTCTGGGAGTCT	3240
iy	3241	TCCATAAGTTTTGCATGAGGACAAACAAACAGGATTAATACTAGTGTGGTTCCTTCAGCC	3300
ib	3241	TCCATAAGTTTTGCATGAGGACAAACAAACAGGATTAATACTAGTGTGGTTCCTTCAGCC	3300
iy	3301	CTCTTAAAGCATAGGGCTTAGCCCTGCAGGCTTCCTTGGGCTTTCTCTGTGTGTAGTTTT	3360
ib	3301	CTCTTAAAGCATAGGGCTTAGCCCTGCAGGCTTCCTTGGGCTTTCTCTGTGTGTAGTTTT	3360
iy	3361	TGTAACACATATACATCTGTTTAAATCCAGTGTCCATGGAACCTTCCACATGCCGTG	3420
ib	3361	TGTAACACATATACATCTGTTTAAATCCAGTGTCCATGGAACCTTCCACATGCCGTG	3420
iy	3421	ACTCTGACATATACGATTTTGAAGACAGGGTTCTCTCTGCCCTGTCTAAACAGCCCAAGT	3480

Db	3421	ACTCTGGACATATACAGTCTTTTGGAAAGCAGGGTTCCTCTGCTCTGCTAAACAAGCCCAAGT	3480
Qy	3481	GGACCAAGTCGAATGCTTTCCTTTACACTATATGTTTTTAAGTAGTCAAACTTCAAGAA	3540
Db	3481	GGACCAAGTCGAATGCTTTCCTTTACACTATATGTTTTTAAGTAGTCAAACTTCAAGAA	3540
Qy	3541	CAATCTAAACAAGTTCTGTGTGCATATGTTTGTGAACTTTGTAATTTGTAATTTAGTAGGC	3600
Db	3541	CAATCTAAACAAGTTCTGTGTGCATATGTTTGTGAACTTTGTAATTTGTAATTTAGTAGGC	3600
Qy	3601	TTCTATATTGCATTTAACTTGTTTTTGTAACTCCTGATCTTCTTTTCGGATACATTG	3660
Db	3601	TTCTATATTGCATTTAACTTGTTTTTGTAACTCCTGATCTTCTTTTCGGATACATTG	3660
Qy	3661	ATGAATAAGAAATTTAAAGTGAAAAA	3693
Db	3661	ATGAATAAGAAATTTAAAGTGAAAAA	3693
RESULT 2			
US-10-084-817-20			
; Sequence 20, Application US/10084817			
; Publication No. US20030119009A1			
; GENERAL INFORMATION:			
; APPLICANT: Susan Stuart			
; APPLICANT: Jed G. Nuchtern			
; APPLICANT: Sharon E. Pion			
; APPLICANT: Jason M. Shohet			
; TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION			
; FILE REFERENCE: PA-0046 US			
; CURRENT APPLICATION NUMBER: US/10/084,817			
; CURRENT FILING DATE: 2002-02-25			
; PRIOR APPLICATION NUMBER: 60/270,784			
; PRIOR FILING DATE: 2001-02-23			
; NUMBER OF SEQ ID NOS: 365			
; SOFTWARE: PERL Program			
; SEQ ID NO 20			
; LENGTH: 3693			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: misc feature			
; OTHER INFORMATION: Incyte ID No. US20030119009A1 1505038CB1			
US-10-084-817-20			
Query Match 100.0%; Score 3693; DB 14; Length 3693;			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 3693; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	CGATCGGGGACTCCACCTCCGCTTACAGCTCGCTCGCGCGTCTCTGCGCGCGGCGCCCC	60
Db	1	CGATCGGGGACTCCACCTCCGCTTACAGCTCGCTCGCGCGTCTCTGCGCGCGGCGCCCC	60
Qy	61	AGGAGACTGTGACACAGACCAAGATGTGGAAACGCTGCTCGCGCTCGCGCTCGCGCTGCT	120
Db	61	AGGAGACTGTGACACAGACCAAGATGTGGAAACGCTGCTCGCGCTCGCGCTCGCGCTGCT	120
Qy	121	GCGGTGCGCTGGGTCCGCGCGGAGAGAGCTAAGAGCAAAATCCAAAGATCTGTGCCAA	180
Db	121	GCGGTGCGCTGGGTCCGCGCGGAGAGAGCTAAGAGCAAAATCCAAAGATCTGTGCCAA	180
Qy	181	TGTGTTTTGTGGAGCCGCGCGGGAATGTGCAGTCAAGAGAAAGGGGAACCACTGTCT	240
Db	181	TGTGTTTTGTGGAGCCGCGCGGGAATGTGCAGTCAAGAGAAAGGGGAACCACTGTCT	240
Qy	241	CTGCATTGAGCAATGCAAACTCAAGAGGCGCTGTGTGGCAGTAAATGCGAAGACCTA	300
Db	241	CTGCATTGAGCAATGCAAACTCAAGAGGCGCTGTGTGGCAGTAAATGCGAAGACCTA	300
Qy	301	CCTCAACCACTGTGAATGCAATGCAATGCTGCTCACTGGATCCAAATCCAGTTGA	360
Db	301	CCTCAACCACTGTGAATGCAATGCAATGCTGCTCACTGGATCCAAATCCAGTTGA	360

```

; NAME/KEY: misc_feature
; OTHER INFORMATION: Incvte ID No. US20030119009A1 1505038CB1

```

Query Match	100.0%;	Score 3693;	DB 14;	Length 3693;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 3693;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	CGATCGGGGACTCCCACTTCGGCTTACAGCTCGCTCGCGCGCTCTGCTCCCGCGCGCCCC	60	
DB	1	CGATCGGGGACTCCCACTTCGGCTTACAGCTCGCTCGCGCGCTCTGCTCCCGCGCGCCCC	60	
QY	61	AGGAGACTCTGACACAGACACGATGTGGAAAAGCTGGCTCGCGCTCGCGCTCGCGCTGGT	120	
DB	61	AGGAGACTCTGACACAGACACGATGTGGAAAAGCTGGCTCGCGCTCGCGCTCGCGCTGGT	120	
QY	121	GGCGGTGCTGTGGTTCGCGCCGAGGAAGCTAAGAGCAAAATCCAGATCTGTGCCAA	180	
DB	121	GGCGGTGCTGTGGTTCGCGCCGAGGAAGCTAAGAGCAAAATCCAGATCTGTGCCAA	180	
QY	181	TGTGTTTTGTGGACCGCGCGGGAATGTGCAGTCACAGAGAAAGGGGAACCCACCTGTCT	240	
DB	181	TGTGTTTTGTGGACCGCGCGGGAATGTGCAGTCACAGAGAAAGGGGAACCCACCTGTCT	240	
QY	241	CTGCATTGAGCAATGCAAACTCACAGAGGCCCTGTGTGTGGCAGTAAATGGCAAGACCTTA	300	
DB	241	CTGCATTGAGCAATGCAAACTCACAGAGGCCCTGTGTGTGGCAGTAAATGGCAAGACCTTA	300	
QY	301	CCTCAACCACTGTGAATCGATCGAGATGGCTGCCTCACTGGATCCAAATCCAGGTTGA	360	
DB	301	CCTCAACCACTGTGAATCGATCGAGATGGCTGCCTCACTGGATCCAAATCCAGGTTGA	360	

QY 361 TTACGATGGACACTGCAAGAGAGAAATCCGTAAAGTCCATCTGTCCAGCCAGCTGTGTTG 420
 Db 361 TTACGATGGACACTGCAAGAGAGAAATCCGTAAAGTCCATCTGTCCAGCCAGCTGTGTTG 420
 QY 421 CTATCAGTCCCAACCGTGTAGTCCGACGTCGCATCATCAGTGGCTGGAAGCTGAGAT 480
 Db 421 CTATCAGTCCCAACCGTGTAGTCCGACGTCGCATCATCAGTGGCTGGAAGCTGAGAT 480
 QY 481 CATCTCAGATGGCTGGTTCTCTAAGGCGACCACTACAGTGAATCCTAGACAAGTATTT 540
 Db 481 CATCTCAGATGGCTGGTTCTCTAAGGCGACCACTACAGTGAATCCTAGACAAGTATTT 540
 QY 541 TAAGAACTTTGATTAATGGTGATCTCCGCTGGACTCCAGTGAATCCTGGAAGTTGTGGA 600
 Db 541 TAAGAACTTTGATTAATGGTGATCTCCGCTGGACTCCAGTGAATCCTGGAAGTTGTGGA 600
 QY 601 ACAGAACTGAACTGCCATCAATATTAACAAGTATCCAGACAGAGAGAACCAAGTTGCT 660
 Db 601 ACAGAACTGAACTGCCATCAATATTAACAAGTATCCAGACAGAGAGAACCAAGTTGCT 660
 QY 661 TAGGGACTCTGTGTGATCTCTCATTTGAATCTCTGTGATGAATCTGATTTGGAACCT 720
 Db 661 TAGGGACTCTGTGTGATCTCTCATTTGAATCTCTGTGATGAATCTGATTTGGAACCT 720
 QY 721 CAGCTTCCAAGAGTTTCTCAAGTGCCTCAACCCATCTTCAACCCATCTTCAACCCCTCCTGAGAAGAGTG 780
 Db 721 CAGCTTCCAAGAGTTTCTCAAGTGCCTCAACCCATCTTCAACCCCTCCTGAGAAGAGTG 780
 QY 781 TGCCCTGGAGGATGAACGATGATGAGTGGAGTGCAGACGAGTGCATCTGTAAACCGCTG 840
 Db 781 TGCCCTGGAGGATGAACGATGATGAGTGGAGTGCAGACGAGTGCATCTGTAAACCGCTG 840
 QY 841 TGTCTGTCCCTGTGGAAATTTGGGTCTGTACAGCCATGACCTGTGACGGAAAGAAATCAGAA 900
 Db 841 TGTCTGTCCCTGTGGAAATTTGGGTCTGTACAGCCATGACCTGTGACGGAAAGAAATCAGAA 900
 QY 901 GGGGGCCAGACCCAGACAGAGAGAGAGATGACACAGATATGTCAGAGCTCCAAAGCA 960
 Db 901 GGGGGCCAGACCCAGACAGAGAGAGAGATGACACAGATATGTCAGAGCTCCAAAGCA 960
 QY 961 TCAGAAACAGCTGAAAGACCAAGAGAGTGCAGACCAAGAGATCTAATGAGAGGACAC 1020
 Db 961 TCAGAAACAGCTGAAAGACCAAGAGAGTGCAGACCAAGAGATCTAATGAGAGGACAC 1020
 QY 1021 AGACAGTGTCTGGATCCAGATCTTCTCACTTCCAGCTGAGTTCAGTATACACAAG 1080
 Db 1021 AGACAGTGTCTGGATCCAGATCTTCTCACTTCCAGCTGAGTTCAGTATACACAAG 1080
 QY 1081 TGTCTGTACAGTCCCAATACACAGATATTTGCTTATAGCAATCAGTATTTTGT 1140
 Db 1081 TGTCTGTACAGTCCCAATACACAGATATTTGCTTATAGCAATCAGTATTTTGT 1140
 QY 1141 TTAATTTGTTTTCGAATAAGGATATGAAGTGGCTGAGTGAAGAGGAAAGGCGCACAGCC 1200
 Db 1141 TTAATTTGTTTTCGAATAAGGATATGAAGTGGCTGAGTGAAGAGGAAAGGCGCACAGCC 1200
 QY 1201 TTCAATTTAGAGTGTCTTAAAGAACTCTAATGGTGTCTGGGGCTGGAGGCTAGT 1260
 Db 1201 TTCAATTTAGAGTGTCTTAAAGAACTCTAATGGTGTCTGGGGCTGGAGGCTAGT 1260
 QY 1261 AAGGAACTGTGATCAGTATGAAGAGGAAACAGACCCAAATCTGAACTCTTTTGTGTTT 1320
 Db 1261 AAGGAACTGTGATCAGTATGAAGAGGAAACAGACCCAAATCTGAACTCTTTTGTGTTT 1320
 QY 1321 ACTGATCTGTGACAGAGCTGAGGAGTGCACAGATGCCAGAGAGAACTTAGCAGGGT 1380
 Db 1321 ACTGATCTGTGACAGAGCTGAGGAGTGCACAGATGCCAGAGAGAACTTAGCAGGGT 1380
 QY 1381 GTCCCGGAGGAGAGGTTTGGGAAGCTCCAGGAGAGGAACGCTCTCTGCTTCCAGGCTC 1440
 Db 1381 GTCCCGGAGGAGAGGTTTGGGAAGCTCCAGGAGAGGAACGCTCTCTGCTTCCAGGCTC 1440
 QY 1441 TTTCCATTTGCCGCTCAGCATGACAGACCTCCAGCATCCACCGCATCTTTGTGTCCTCAATAC 1500

Db 1441 TTTCCATTTGCCGCTCAGCATGACAGACCTCCAGCATCCACCGCATCTTTGGTCCCAATAC 1500
 QY 1501 TGCCTCTAGATACATAGCCATAGTCTAGTTAAACCCAGTGTCCCTCAGACTTGGATGGAG 1560
 Db 1501 TGCCTCTAGATACATAGCCATAGTCTAGTTAAACCCAGTGTCCCTCAGACTTGGATGGAG 1560
 QY 1561 TTTCTGGAGGGGTACACCCAAATGATGAGATCTTGTATATCTTTGAGCCCTTAGCCAC 1620
 Db 1561 TTTCTGGAGGGGTACACCCAAATGATGAGATCTTGTATATCTTTGAGCCCTTAGCCAC 1620
 QY 1621 CTAAACCAAAATTTTAAATACTTTTACCAAGGTGCTATTTCTCTGTAAACACACTTTT 1680
 Db 1621 CTAAACCAAAATTTTAAATACTTTTACCAAGGTGCTATTTCTCTGTAAACACACTTTT 1680
 QY 1681 TTTGGCAAGTTGACTTTTATCTTCAATTTATTAATCATATATATTTTAAATATTTT 1740
 Db 1681 TTTGGCAAGTTGACTTTTATCTTCAATTTATTAATCATATATATTTTAAATATTTT 1740
 QY 1741 ATTTTCTTGACTAGTATTAAGCTTTTGTATTTATTTTTCAGTAGTCCACCACTTCATA 1800
 Db 1741 ATTTTCTTGACTAGTATTAAGCTTTTGTATTTATTTTTCAGTAGTCCACCACTTCATA 1800
 QY 1801 GGTGAAAGAGTTTGGGTTCTTCTGTGTGAGGGGCTGAAATAACCCAGATGCTCCAC 1860
 Db 1801 GGTGAAAGAGTTTGGGTTCTTCTGTGTGAGGGGCTGAAATAACCCAGATGCTCCAC 1860
 QY 1861 CCTGCCACATATAGATGACGCCATAGTGTGCCCCCTAGCTTCCAGCAGTCCACTATC 1920
 Db 1861 CCTGCCACATATAGATGACGCCATAGTGTGCCCCCTAGCTTCCAGCAGTCCACTATC 1920
 QY 1921 TGCCAGAGGAGCAAGGTGCCCTTAGACCTAAGCCAGGGGAAAGAGCATCTTCAATAAAAA 1980
 Db 1921 TGCCAGAGGAGCAAGGTGCCCTTAGACCTAAGCCAGGGGAAAGAGCATCTTCAATAAAAA 1980
 QY 1981 CTTTCAAGATCCAAACATTAATTTGTTTTTATTTTCTGAGAGTTCAGGCAATCAGT 2040
 Db 1981 CTTTCAAGATCCAAACATTAATTTGTTTTTATTTCTGAGAGTTCAGGCAATCAGT 2040
 QY 2041 ATTCCCAAGGATGGGCAAGGCGCAGCCAGCGGCTTAGGATATCCAGCCTACCAAT 2100
 Db 2041 ATTCCCAAGGATGGGCAAGGCGCAGCCAGCGGCTTAGGATATCCAGCCTACCAAT 2100
 QY 2101 ATGCTCATTCGACTAATCTAGAGGGTGTGTTGGCCCTGTCTCTCTTTTCTGACCTC 2160
 Db 2101 ATGCTCATTCGACTAATCTAGAGGGTGTGTTGGCCCTGTCTCTCTTTTCTGACCTC 2160
 QY 2161 AGTTTCTCTCAGTGAGCTGGTAAGAAATGCATAACCTTTTGTATTTGATTAAGTTATAAATC 2220
 Db 2161 AGTTTCTCTCAGTGAGCTGGTAAGAAATGCATAACCTTTTGTATTTGATTAAGTTATAAATC 2220
 QY 2221 TGTGTTCTGTATCTGATGCTCAGAGGGGAGATAGTTCCTGTGATTTTCTTCTTCTCT 2280
 Db 2221 TGTGTTCTGTATCTGATGCTCAGAGGGGAGATAGTTCCTGTGATTTTCTTCTTCTCT 2280
 QY 2281 ATAGAAATAAATCAAACTTTGTTACTAGAAACAGAAATGTAGATGGCCCAAAACAAAGATG 2340
 Db 2281 ATAGAAATAAATCAAACTTTGTTACTAGAAACAGAAATGTAGATGGCCCAAAACAAAGATG 2340
 QY 2341 ACCAGATTTGATCTCAGCCTGATGACCTCCTACAGTCTGCTATGATATGGAGTCCCTATG 2400
 Db 2341 ACCAGATTTGATCTCAGCCTGATGACCTCCTACAGTCTGCTATGATATGGAGTCCCTATG 2400
 QY 2401 GGTAAAGCAGGAGAGAGTGGGAAAGAGAACCCACCCACTCTCTCTCATATTTTGCATTT 2460
 Db 2401 GGTAAAGCAGGAGAGAGTGGGAAAGAGAACCCACCCACTCTCTCTCATATTTTGCATTT 2460
 QY 2461 CATGTTTAACTCCGGCTGGAATAAGAAAGCATTCCTTTAGAGATGAGGATTAAGAAAG 2520
 Db 2461 CATGTTTAACTCCGGCTGGAATAAGAAAGCATTCCTTTAGAGATGAGGATTAAGAAAG 2520
 QY 2521 TTTTCAGATTCACAGGGGAGAAATGGAGATTTAATCCCTAAACCTGTGACTTGGGAG 2580

2y	552	ATAATGGTGATTCCTCGCCTGGACTCCAGTGAATTCCTGAAGTTTGTGGAAACAAGATGAAA	611
2b	561	ATAATGGTGATTCCTCGCCTGGACTCCAGTGAATTCCTGAAGTTTGTGGAAACAAGATGAAA	620
2y	612	CTGCCATCAATATTACAAAGTATCCAGACACAGAGAAACAAGTTGCTTTAGGGGACTCT	671
2b	621	CTGCCATCAATATTACAAAGTATCCAGACACAGAGAAACAAGTTGCTTTAGGGGACTCT	680
2y	672	GTGTTGATGCTCTCAITGAACTGCTGTATGAAATATCTGATGGAAATCAGCTTCCAAAG	731
2b	681	GTGTTGATGCTCTCAITGAACTGCTGTATGAAATATCTGATGGAAATCAGCTTCCAAAG	740
2y	732	AGTTTCTCAAGTGCCTCAACCCATCTTTCAACCCCTCTCGAGAGAAAGTGTGCCCTGGAG	791
2b	741	AGTTTCTCAAGTGCCTCAACCCATCTTTCAACCCCTCTCGAGAGAAAGTGTGCCCTGGAG	800
2y	792	ATGAAACGATATGCAGATGGAGCTGAGACCGAGGTGGACTGTAAACCGCTGTGTCTGTGCCT	851
2b	801	ATGAAACGATATGCAGATGGAGCTGAGACCGAGGTGGACTGTAAACCGCTGTGTCTGTGCCT	860
2y	852	GTGGAATTTGGGTCTGTACAGCATGACCTGTGACGGMAAGATCAGAAGGGGGCCACGA	911
2b	861	GTGGAATTTGGGTCTGTACAGCATGACCTGTGACGGMAAGATCAGAAGGGGGCCACGA	920
2y	912	CCCAGACAGAGGAGGAGATCACCAGATATCTCCAGAGACTCCAAAAGCATCAGGAAACAG	971
2b	921	CCCAGACAGAGGAGGAGATCACCAGATATCTCCAGAGACTCCAAAAGCATCAGGAAACAG	980
2y	972	CTGAAAGACCAAGAGAGTGAGACCAAGAGATCTAATGAGAGAGCACAGACCATGTTC	1031
2b	981	CTGAAAGACCAAGAGAGTGAGACCAAGAGATCTAATGAGAGAGCACAGACCATGTTC	1040
Qy	1032	TGATATCCAGCATCTTCTCCACTTCAGCGCTGAGTTCAGTATACAAAGTGTCTGTCTACA	1091
Db	1041	TGATATCCAGCATCTTCTCCACTTCAGCGCTGAGTTCAGTATACAAAGTGTCTGTCTACA	1100
Qy	1092	GTGCCAAATCACAGTATTTGTTATATAGCAATAGATTTTATTTTGTATTGTTTTT	1151
Db	1101	GTGCCAAATCACAGTATTTGTTATATAGCAATAGATTTTATTTTGTATTGTTTTT	1160
2y	1152	GCAATAAGAGATATGAAGTGGCTGGCTAGAAAGGAAGGCCACAGCCCTTCATTTCCTAG	1211
2b	1161	GCAATAAGAGATATGAAGTGGCTGGCTAGAAAGGAAGGCCACAGCCCTTCATTTCCTAG	1220
2y	1212	GAGTGTCTTTAAGAGAAACTGTAAATGGTCTCTGGGCTCGAGGCTAGTAAGGAAACTGTC	1271
2b	1221	GAGTGTCTTTAAGAGAAACTGTAAATGGTCTCTGGGCTCGAGGCTAGTAAGGAAACTGTC	1280
Qy	1272	ATCAGGATGAAAGAGAAACAGACCCAAATCTGAACTCTTTTGTAGTTTACTGCATCTGT	1331
2b	1281	ATCAGGATGAAAGAGAAACAGACCCAAATCTGAACTCTTTTGTAGTTTACTGCATCTGT	1340
Qy	1332	CAGCAGGCTGAGGGAGTGACACGATGCCAGAGAGAACTTAGCAGGGTGTCCCCGAGG	1391
2b	1341	CAGCAGGCTGAGGGAGTGACACGATGCCAGAGAGAACTTAGCAGGGTGTCCCCGAGG	1400
Qy	1392	AGAGGTTTGGAAAGCTCCACGGAGAGAAACGCTCTCTGCTTCCAGCCTCTTTTCGATTGCC	1451
2b	1401	AGAGGTTTGGAAAGCTCCACGGAGAGAAACGCTCTCTGCTTCCAGCCTCTTTTCGATTGCC	1460
Qy	1452	GTACGATGACAGACCTCCAGCATCTCTGGTCCCAATAACTGCGCTCTAGAT	1511
2b	1461	GTACGATGACAGACCTCCAGCATCTCTGGTCCCAATAACTGCGCTCTAGAT	1520
Qy	1512	ACATAGCCATATCTGTAGTTAAACCAAGTGTCCCTCAGACTTTGGATGAGAGTTTCTGGAGG	1571
2b	1521	ACATAGCCATATCTGTAGTTAAACCAAGTGTCCCTCAGACTTTGGATGAGAGTTTCTGGAGG	1580
Qy	1572	GTACACCCAAATGATGCAGATACCTTTGATATCTTTGAGCCCTTAGCCACTTAACCAATTT	1631
2b	1581	GTACACCCAAATGATGCAGATACCTTTGATATCTTTGAGCCCTTAGCCACTTAACCAATTT	1640
Qy	1632	TTAAAAATACCTTTTTCACAAAGGTGCTATTTCTCTGTAAAAACACTTTTTTTTGGCAAGTT	1691

[illegible]

RESIST 4

```

US-09-960-706-941
; Sequence 941, Application US/09960706
; Publication No. US20030134280A1
; GENERAL INFORMATION:
; APPLICANT: Munger, William E.
; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benig
; TITLE OF INVENTION: Gene Expression Profiles
; FILE REFERENCE: 44921-5029-01US
; CURRENT APPLICATION NUMBER: US/09/960,706
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 60/223,323
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: 09/873,319
; PRIOR FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 941
; LENGTH: 1987
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20030134280A1 U06863
US-09-960-706-941

```

Query Match	52.5%;	Score 1940;	DB 10;	Length 1987;
Best Local Similarity	99.7%;	Pred. No. 0;		
Matches 1954;	Conservative 0;	Mismatches 5;	Indels 1;	Gaps 17;
QY	12	CTCCACACTCCGCTTACAGCTCGCTCGCCGCGCTCTCTCCCGCGCCGCCACAGAGACCTGG	71	
DB	21	CTCCACCTCCGCTTACAGCTCGCTCGCCGCGCTCTCTCCCGCGCCGCCACAGAGACCTGG	80	
QY	72	ACCAGACCACGATGTGGAACGCTGGCTCGCGCTCGCGCTCGCGCTGGTGCAGCTCGGCT	131	
DB	81	ACCAGACCACGATGTGGAACGCTGGCTCGCGCTCGCGCTCGCGCTGGTGCAGCTCGGCT	140	
QY	132	GGGTCCGCGCGGAGGAAGAGCTAAGGAGCAATCCAAGATCTGTGCCAATGTGTTTTGTG	191	
DB	141	GGGTCCGCGCGGAGGAAGAGCTAAGGAGCAATCCAAGATCTGTGCCAATGTGTTTTGTG	200	
QY	152	GAGCGCGCGGGAATGTGCAGTCTACAGAGAAAGGGGAACCCACCTGTCTCTGCATTGAGC	251	
DB	201	GAGCGCGCGGGAATGTGCAGTCTACAGAGAAAGGGGAACCCACCTGTCTCTGCATTGAGC	260	
QY	252	AATGCAAACTTCAAGAGGCGCTGTGTGTGGCAGTAAATGGCAAGACCTTACCAACCT	311	
DB	261	AATGCAAACTTCAAGAGGCGCTGTGTGTGGCAGTAAATGGCAAGACCTTACCAACCT	320	

QY	312	GTGAACGCAATCGAGATAGCGTGCCTCACTGGATCCAAAATCCAGAGTTGATTAAGATGAGC	371
Db	321	GTGAATCTGCAATCGAGATGCTGCTCACTGGAATCCAAAATCCAGAGTTGATTAAGATGAGC	380
QY	372	ACTGCAAAAGAAAGAAATCCGTAACTGCAATCTGCAGAGCCCAAGTTGTTGCTATCACTCCA	431
Db	381	ACTGCAAAAGAAAGAAATCCGTAAAGTCCATCTCCAGACCAGATGTTGGCTATCACTGCCA	440
QY	432	ACCGTATGAGCTCCGACGTCCGATCAATCCAGTGGGCTGGAAGCTGAGATCATCTCAGATG	491
Db	441	ACCGTATGAGCTCCGACGTCCGATCATCTCAAGTGGCTGGAAGCTGAGATCATCTCCAGATG	500
QY	492	GCTGGTTCTCTAAAGGCAAGCACTACGTGAATCCTTAGCAAGTATTTTAAGAATTG	551
Db	501	GCTGGTTCTCTAAAGGCAAGCACTACGTGAATCCTTAGCAAGTATTTTAAGAATTG	560
QY	552	ATATGAGTGAATTCGCGCTGGAATCCAGTGAATTCCTGAAGTTGTGGAACAGATGAAA	611
Db	561	ATATGAGTGAATTCGCGCTGGAATTCCTGAAGTTGTGGAACAGATGAAA	620
QY	612	CTGCCATCATATTTACAACGTATTCAGACCCAGAGAACAAACAAGTTCCTTAAGGGAATCT	671
Db	621	CTGCCATCATATTTACAACGTATTCAGACCCAGAGAACAAACAAGTTCCTTAAGGGAATCT	680
QY	672	GTTGTAATGCTCTCATTTGAATCTGTCTATGAAAATGCTGATTTGAAAATCAAGTCTCCAG	731
Db	681	GTTGTAATGCTCTCATTTGAATCTGTCTATGAAAATGCTGATTTGAAAATCAAGTCTCCAG	740
QY	732	AGTTTCTCAAGTGGCTCAACCCATCTTTCAACCTCCTGAGAAAATAATGTGCCCCGGAGG	791
Db	741	AGTTTCTCAAGTGGCTCAACCCATCTTTCAACCTCCTGAGAAAATAATGTGCCCCGGAGG	800
QY	792	ATGAAACGTATGCAATGAGAGCTAGACCCAGGTGGAATGTTAAACCGCTGTCTCTGTGCT	851
Db	801	ATGAAACGTATGCAATGAGAGCTAGACCCAGGTGGAATGTTAAACCGCTGTCTCTGTGCT	860
QY	852	GTGGAATTTGGGCTCTGTACAGCCATGACCTGTGACGGAAAATAATCAAGGGGGCCAGA	911
Db	861	GTGGAATTTGGGCTCTGTACAGCCATGACCTGTGACGGAAAATAATCAAGGGGGCCAGA	920
QY	912	CCCAAGACAGAGAGAGATGACCAAGATATGTCCAGAGAGCTCCAAAGATCAGGAAACAG	971
Db	921	CCCAAGACAGAGAGAGATGACCAAGATATGTCCAGAGAGCTCCAAAGATCAGGAAACAG	980
QY	972	CTGAAAAGACCAACAGATGAGCAACCAAGATCTTAATGAGAGGCAACAGACAGATGC	1031
Db	981	CTGAAAAGACCAACAGATGAGCAACCAAGATCTTAATGAGAGGCAACAGACAGATGC	1040
QY	1032	TGGAATCCAGATCTTCTCAACTTCAGGCTGAGTTCAATACAAAGTGTGCTACACA	1091
Db	1041	TGGAATCCAGATCTTCTCAACTTCAGGCTGAGTTCAATACAAAGTGTGCTACACA	1100
QY	1092	GTCCGCCAAATCAACCAAGTATTTGCTTAATTAACAAATGATTTTATTTGTTATTTGTTT	1151
Db	1101	GTCCGCCAAATCAACCAAGTATTTGCTTAATTAACAAATGATTTTATTTGTTATTTGTTT	1160
QY	1152	GCAATTAAGATATGAAAGGTGAGCTGGCTAGAAAGGAAAGGGCCACAGCTTCATTTCTAG	1211
Db	1161	GCAATTAAGATATGAAAGGTGAGCTGGCTAGAAAGGAAAGGGCCACAGCTTCATTTCTAG	1220
QY	1212	GAGTGCTTTAAGAGAAACTGTAAATGAGTGCTGTGGGGCTGAGAGCTAATAGAAACATGC	1271
Db	1221	GAGTGCTTTAAGAGAAACTGTAAATGAGTGCTGTGGGGCTGAGAGCTAATAGAAACATGC	1280
QY	1272	ATCACGATTTGAAAGAGAACAGACCCAAATCTGAACCTCTTTTGAGTTTACTGCATCTGT	1331
Db	1281	ATCACGATTTGAAAGAGAACAGACCCAAATCTGAACCTCTTTTGAGTTTACTGCATCTGT	1340
QY	1332	CAGCAGGCTGCAAGGAGTGCACACATGCAAGAGAGAACTTAAGCAGGCTGTCCCGAGAG	1391
Db	1341	CAGCAGGCTGCAAGGAGTGCACACATGCAAGAGAGAACTTAAGCAGGCTGTCCCGAGAG	1400
QY	1392	AGAGGTTTGGAAAGCTCCACGAGAGAAAGCTCTCTGCTTCCAGCTCTTTCCATTGGC	1451

Db	1401	AGAGGTTTGGGAAGCTCCAGAGAGAGGAAGCGTCTCGTTCAGCGCTCTTCCATGAGCC	1460
QY	1452	GTCAGCATGACAGAGCCTCCAGCATCCAGCGATCTCTTGATGCCAATAAGCCGCTTGAT	1511
Db	1461	GTCAGCATGACAGAGCCTCCAGCATCCAGCGATCTCTTGATGCCAATAAGCCGCTTGAT	1520
QY	1512	ACATAGGCCTACTGCTAGTTAAACCCAGTGTCCTCCAGACTTGATGGAGTTCTGGAGG	1571
Db	1521	ACATAGGCCTACTGCTAGTTAAACCCAGTGTCCTCCAGACTTGATGGAGTTCTGGAGG	1580
QY	1572	GTAACACCCAAATGATGCAGTACTGTATCTTGATCTTGAGCCCTTGAGCGACTAACCAATT	1631
Db	1581	GTAACACCCAAATGATGCAGTACTGTATCTTGAGCCCTTGAGCGACTAACCAATT	1640
QY	1632	TTAATAATCTTTTATACCAAGGAGCATTTCTGTAAACAACATTTTCTTTGGCAAGTT	1691
Db	1641	TTAATAATCTTTTATACCAAGGAGCATTTCTGTAAACAAC-TTTTCTGGCAAGTT	1699
QY	1692	GACTTTATCTTCATTTATTATCATATATATTGTGTTTTTAATATTATTTCTTGAC	1751
Db	1700	GACTTTATCTTCATTTATTATCATATATATTGTGTTTTTAATATTATTTCTTGAC	1759
QY	1752	TAGTATTAAAGCTTTTGATTTAATTTTCAGTGTCCACACACTCAATAGGTGAAGAG	1811
Db	1760	TAGTATTAAAGCTTTTGATTTAATTTTCAGTGTCCACACACTCAATAGGTGAAGAG	1819
QY	1812	TTTGAGGCTTCTTCCTGTGCAGGGGCTGAATAAACCCAGATGCTCCACCCCTGCACATA	1871
Db	1820	TTTGAGGCTTCTTCCTGTGCAGGGGCTGAATAAACCCAGATGCTCCACCCCTGCACATA	1879
QY	1872	CTAGATGACGCCATAGTTGGCCCCCTAGCTTCAGAGTCCACTATCTGCCAGAGAG	1933
Db	1880	CTAGATGACGCCATAGTTGGCCCCCTAGCTTCAGAGTCCACTATCTGCCAGAGAG	1939
QY	1932	CAAGGAGGCTTAGACTTAAGCCAGGAGGAAGAAGCATCTT	1971
Db	1940	CAAGGAGGCTTAGACTTAAGCCAGGAGGAAGAAGCATCTT	1979

```

RESULT 5
; US-09-873-319-617
; Sequence 617, Application US/09873319A
; Publication No. US20030134324A1
; GENERAL INFORMATION:
; APPLICANT: Murger, William E.
; APPLICANT: Kulkarni, Prakash
; APPLICANT: Getzenberg, Robert H.
; APPLICANT: Waga, Iwao
; APPLICANT: Yamamoto, Jun
; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic
; TITLE OF INVENTION: Hypertasia Using Gene Expression Profiles
; FILE REFERENCE: 44921-5029-US
; CURRENT APPLICATION NUMBER: US/09/873.319A
; CURRENT FILING DATE: 2001-06-05
; EARLIER APPLICATION NUMBER: US 60/223.323
; EARLIER FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 755
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 617
; LENGTH: 1987
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20030134324A1 U06863
US-09-873-319-617

Query Match          52.5%; Score 1940; DB 10; Length 1987;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1954; Conservative 0; Mismatches 5; Indels 1; Gaps 1

12 CTCACACTCGGCTAGCAGCTGGTGCAGCCGCGTCTGTGCGCGCCCCAGAGACTCGG 71
|||||

```

Db 21 CTCACACTCGCTTACAGCTCGCTCGCGCGCTCTCTCGCCCGCCCGCCAGGAGACCTCG 80
Qy 72 ACCAGACACGATGTGAAACGCTGGCTCGCGCTCGCGCTCGCGCTGGTGGCGGTGCGCT 131
Db 81 ACCAGACACGATGTGAAACGCTGGCTCGCGCTCGCGCTCGCGCTGGTGGCGGTGCGCT 140
Qy 132 GGGTCCGCGCGCGAGAGAGCTAAGGAGCAAAATCCAGATCTGTGCAATGTGTTTGTG 191
Db 141 GGGTCCGCGCGCGAGAGAGCTAAGGAGCAAAATCCAGATCTGTGCAATGTGTTTGTG 200
Qy 192 GAGCGCGCGGGAATGTGAGTCAAGAGAAAGGGAACCCACCTGTCTGTGCAATGTGAGC 251
Db 201 GAGCGCGCGGGAATGTGAGTCAAGAGAAAGGGAACCCACCTGTCTGTGCAATGTGAGC 260
Qy 252 AATGCAAACTCCACAGAGGCTGTGTGGCAGTAATGCAAGACCTACTCAACCACT 311
Db 261 AATGCAAACTCCACAGAGGCTGTGTGGCAGTAATGCAAGACCTACTCAACCACT 320
Qy 312 GTGAATGTGATGTGAGTCCCTGCTCATCTGATGCAATCCAAATCCAGGTGATGATGAGC 371
Db 321 GTGAATGTGATGTGAGTCCCTGCTCATCTGATGCAATCCAAATCCAGGTGATGATGAGC 380
Qy 372 ACTGCAAGAGAGAAATCGTAAGTCCATCTGCGACCCAGTGTGTTGTATCAGTCCA 431
Db 381 ACTGCAAGAGAGAAATCGTAAGTCCATCTGCGACCCAGTGTGTTGTATCAGTCCA 440
Qy 432 ACCGTGATGACTCCGAGTCCGATCATCTGAGTGGTGGAGGTGAGATCATCTCAGATG 491
Db 441 ACCGTGATGACTCCGAGTCCGATCATCTGAGTGGTGGAGGTGAGATCATCTCAGATG 500
Qy 492 GCTGGTCTCTAAGAGGAGCACTACAGTGAATCTTAGCAAGTATTTTAAAGACTTTG 551
Db 501 GCTGGTCTCTAAGAGGAGCACTACAGTGAATCTTAGCAAGTATTTTAAAGACTTTG 560
Qy 552 ATAATGTGATTTCTCGCTGAGCTCCAGTGAATCTTGAAGTGTGGAACAGATGAAA 611
Db 561 ATAATGTGATTTCTCGCTGAGCTCCAGTGAATCTTGAAGTGTGGAACAGATGAAA 620
Qy 612 CTGCGATCAATATTACAGTATCCAGACAGGAGACACAGTGTGCTTAGGGGACTCT 671
Db 621 CTGCGATCAATATTACAGTATCCAGACAGGAGAACACAGTGTGCTTAGGGGACTCT 680
Qy 672 GTGTGATGTCTCTCATTAAGTCTGATGAAAATGCTGATGAAACTCAGCTTCCAAG 731
Db 681 GTGTGATGTCTCTCATTAAGTCTGATGAAAATGCTGATGAAACTCAGCTTCCAAG 740
Qy 732 AGTTTCTCAAGTGCTCAACCCATCTTCAACCTCTGAGAGAAAGTGTGCTTGGAGG 791
Db 741 AGTTTCTCAAGTGCTCAACCCATCTTCAACCTCTGAGAGAAAGTGTGCTTGGAGG 800
Qy 792 ATGAAAAGTATGACAGTGGAGCTGAGACCGAGTGGACTGTAAACCGCTGTGTGTGCT 851
Db 801 ATGAAAAGTATGACAGTGGAGCTGAGACCGAGTGGACTGTAAACCGCTGTGTGTGCT 860
Qy 852 GTGGAATTTGGGTCTGTACAGCCATGACCTGTGACGGAAGATCAGAGGGGGCCGAGA 911
Db 861 GTGGAATTTGGGTCTGTACAGCCATGACCTGTGACGGAAGATCAGAGGGGGCCGAGA 920
Qy 912 CCCAGACAGAGAGAGATGACAGATATGTCAGGAGCTCCAAAGACATCAGAAACAG 971
Db 921 CCCAGACAGAGAGAGATGACAGATATGTCAGGAGCTCCAAAGACATCAGAAACAG 980
Qy 972 CTGAAAAGACCAAGAGAGTGAAGCCAAAGAGATCTAATGAGGAGGACACAGACAGTGTG 1031
Db 981 CTGAAAAGACCAAGAGAGTGAAGCCAAAGAGATCTAATGAGGAGGACACAGACAGTGTG 1040
Qy 1032 TGGATCCAGCATCTTCTCCACTTCCAGCGCTGAGTTCAGTATACACAAGTGTGCTTACA 1091
Db 1041 TGGATCCAGCATCTTCTCCACTTCCAGCGCTGAGTTCAGTATACACAAGTGTGCTTACA 1100
Qy 1092 CTCGCAAAATCACCAGTATTTGCTTATATAGCAATGATTTATTTTGTATTTTGT 1151
Db 1101 CTCGCAAAATCACCAGTATTTGCTTATATAGCAATGATTTATTTTGTATTTTGT 1160

Qy 1152 GCATTAAGGATATGAGTGGCTGGCTAGGAAGGAGGCGCCACAGCCTTCATTTCTAG 1211
Db 1161 GCATTAAGGATATGAGTGGCTGGCTAGGAAGGAGGCGCCACAGCCTTCATTTCTAG 1220
Qy 1212 GAGTCTTTAAGAGAAACTGTAAATGTGTCTCTGGGCTGGAGCTAGTAAGGAACTGC 1271
Db 1221 GAGTCTTTAAGAGAAACTGTAAATGTGTCTCTGGGCTGGAGCTAGTAAGGAACTGC 1280
Qy 1272 ATCAGATTGAAGAGGAGACAGCCAAATCTGAACCTCTTTTCTGAGTTTACTGATCTGT 1331
Db 1281 ATCAGATTGAAGAGGAGACAGCCAAATCTGAACCTCTTTTCTGAGTTTACTGATCTGT 1340
Qy 1332 CAGCAGGCTGAGGAGTGCACAGATGCCAGAGAGAACTTTAGCAGGCTGTCCCGGAGG 1391
Db 1341 CAGCAGGCTGAGGAGTGCACAGATGCCAGAGAGAACTTTAGCAGGCTGTCCCGGAGG 1400
Qy 1392 AGAGTTTGGGAGCTCCACGAGAGGAGACCTCTCTGCTTCCAGCCTCTTTTCCATTTGCC 1451
Db 1401 AGAGTTTGGGAGCTCCACGAGAGGAGACCTCTCTGCTTCCAGCCTCTTTTCCATTTGCC 1460
Qy 1452 GTACAGTGAAGAGACCTCCAGCATCCAGCATCTCTTGGTCCCAATTAACCTCTCTAGAT 1511
Db 1461 GTACAGTGAAGAGACCTCCAGCATCCAGCATCTCTTGGTCCCAATTAACCTCTCTAGAT 1520
Qy 1512 ACATAGCCATCTGCTAGTTAACCCAGTGTCCCTCAGACTTGGATGGAGTTTCTGGGAGG 1571
Db 1521 ACATAGCCATCTGCTAGTTAACCCAGTGTCCCTCAGACTTGGATGGAGTTTCTGGGAGG 1580
Qy 1572 GTACACCCAAATGATGAGATACCTTGTATATCTTTGAGCCCTTAGCGACCTTAAACCAATT 1631
Db 1581 GTACACCCAAATGATGAGATACCTTGTATATCTTTGAGCCCTTAGCGACCTTAAACCAATT 1640
Qy 1632 TTAAAAATCTTTTACCAAGGCTGATTTCTCTGTAAACACATTTTGTGGCAAGTT 1691
Db 1641 TTAAAAATCTTTTACCAAGGCTGATTTCTCTGTAAACAC- TTTTGTGGCAAGTT 1699
Qy 1692 GACTTTATCTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTCTTGAC 1751
Db 1700 GACTTTATCTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTTCTTGAC 1759
Qy 1752 TAGTATTAAAGCTTTTGAATTAATTTTTCAGTGTGCCACCACTTTCATAGTGGAGGAG 1811
Db 1760 TAGTATTAAAGCTTTTGAATTAATTTTTCAGTGTGCCACCACTTTCATAGTGGAGGAG 1819
Qy 1812 TTTGGGTTCTTCTGCTGAGGAGGCTGAAATAACCCAGATGCTCCACCTGCCACATA 1871
Db 1820 TTTGGGTTCTTCTGCTGAGGAGGCTGAAATAACCCAGATGCTCCACCTGCCACATA 1879
Qy 1872 CTAGATGAGCCCATAGTTGGCCCCCTTAGCTTCCAGCAGTCCACTATCTGCCAGAGGAG 1931
Db 1880 CTAGATGAGCCCATAGTTGGCCCCCTTAGCTTCCAGCAGTCCACTATCTGCCAGAGGAG 1939
Qy 1932 CAAGGCTCCCTTAGACCTAAGCCAGGAGGAGCAAGCATCTT 1971
Db 1940 CAAGGCTCCCTTAGACCTAAGCCAGGAGGAGCAAGCATCTT 1979

RESULT 6
US-10-098-841-144/c
; Sequence 144, Application US/10098841
; Publication No. US20020197679A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Xu, Chongjun
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungqing
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhao, Qing A.
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong


```

Db      1621 TTTAACTGTTTTGTAACCTGATCTCTCCCTTTTCGGATATCTATGATGAATAAGAA 1680
Qy      3673 ATTAAGTGA 3682
Db      1681 ATTAAGTGA 1690

RESULT 8
US-10-242-515-3903
; Sequence 3903, Application US/10242515
; Publication No. US20040009488A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005C1
; CURRENT APPLICATION NUMBER: US/10/242,515
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 09/764,877
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,896
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: Patentin ver. 2.0
; SEQ ID NO 3903
; LENGTH: 1755
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-242-515-3903

Query Match      45.4%; Score 1678.4; DB 15; Length 1755;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1682; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy      1993 AAACATTAATTTGTTTTTATTTATTTCTGAGAAAGTTGAGGCAAAATCAGTATTTCCCAAGGAT 2052
Db      1 ANANTTTTTTGTGTTTTTATTTATTTCTGAGAAAGTTGAGGCAAAATCAGTATTTCCCAAGGAT 60

Qy      2053 GCGCAAGAGGCGAGCAAGCAGGGCTTAGGATATCCAGCCCTACAAATATGCTCATTCGA 2112
Db      61 GCGCAAGAGGCGAGCAAGCAGGGCTTAGGATATCCAGCCCTACCAATATGCTCATTCGA 120

Qy      2113 CTAACCTAGAGGGTGAAGTTGGCCCTGCTCTCTTTTCTGGACCTCAGTTTCTCAGT 2172
Db      121 CTAACCTAGAGGGTGAAGTTGGCCCTGCTCTCTTTTCTGGACCTCAGTTTCTCAGT 180

Qy      2173 GAGCTGGTAAGAAATGCACCTAACCTTTGATTTGATTAAGTATAAATTCCTGGTTCGAT 2232

```

```

Db      181 GAGCTGGTAAGAAATGCACATAACCTTTTGATTTGATTAAGTATAAATTCGTGGTTCGAT 240
Qy      2233 CATTGGTCCAGAGGGGAGATAGGTTCCCTGTGATTTTCCCTTCTCTCTATAGAAATAAATG 2292
Db      241 CATTGGTCCAGAGGGGAGATAGGTTCCCTGTGATTTTCCCTTCTCTCTATAGAAATAAATG 300
Qy      2293 AAATCTTGTTACTAGAAACAAGAAATGTCAGATGGCCAAAACCAAGATACCAGATTTGAT 2352
Db      301 AAATCTTGTTACTAGAAACAAGAAATGTCAGATGGCCAAAACCAAGATACCAGATTTGAT 360
Qy      2353 CTCAGCTGATGACCCCTACAGTCTGTGATGATGAGTCCCTCATGGTAAAGCAGGA 2412
Db      361 CTCAGCTGATGACCCCTACAGTCTGTGATGATGAGTCCCTCATGGTAAAGCAGGA 420
Qy      2413 AGAGAGTGGGAAAGAAAGAACCCCACTCTGTCTTTCATATTGCAATTCATGTTTAACT 2472
Db      421 AGAGAGTGGGAAAGAAAGAACCCCACTCTGTCTTTCATATTGCAATTCATGTTTAACT 480
Qy      2473 CCGGCTGGAAATAGAAAGCAATTCCTTAGAGATGAGGATAAAGAAAGTTTCAGATTCAA 2532
Db      481 CCGGCTGGAAATAGAAAGCAATTCCTTAGAGATGAGGATAAAGAAAGTTTCAGATTCAA 540
Qy      2533 CAGGGGGAAGAAATGGAGATTTAATCCTTAAACCTGTGACTTGGGAGGTCAGTCAATTA 2592
Db      541 CAGGGGGAAGAAATGGAGATTTAATCCTTAAACCTGTGACTTGGGAGGTCAGTCAATTA 600
Qy      2593 CAGTTAGTCTGTGTCTTTCGACTTCTGTGATTTAAACCCCACTCACTACCCCTGTTCA 2652
Db      601 CAGTTAGTCTGTGTCTTTCGACTTCTGTGATTTAAACCCCACTCACTACCCCTGTTCA 660
Qy      2653 GATGCATTTGGAATACCAAGATTTAAATCCTTTGACATAAGATTCATTTGCAGAAAGCAG 2712
Db      661 GATGCATTTGGAATACCAAGATTTAAATCCTTTGACATAAGATTCATTTGCAGAAAGCAG 720
Qy      2713 ATTAAGACCATCAGAGGAAATTTAGGTTGTAATGCACAGGCAACTGTGAGAACT 2772
Db      721 ATTAAGACCATCAGAGGAAATTTAGGTTGTAATGCACAGGCAACTGTGAGAACT 780
Qy      2773 GTTGTCGCAAAAATAGAAATTCCTTCTAGTTTCTTGTGTTTCTCATTTGAAAGGAGAAAT 2832
Db      781 GTTGTCGCAAAAATAGAAATTCCTTCTAGTTTCTTGTGTTTCTCATTTGAAAGGAGAAAT 840
Qy      2833 CCACTTGTTTAGCATTTCAAGCTTTTATGATTCATCCCATCTTAAACCTCTTCAACT 2892
Db      841 CCACTTGTTTAGCATTTCAAGCTTTTATGATTCATCCCATCTTAAACCTCTTCAACT 900
Qy      2893 CCACTTGTTTAGCTCTGAAATGCAGCTCCCTGTCCAAAGTCCCTTGGAGAACTCACAGCAGC 2952
Db      901 CCACTTGTTTAGCTCTGAAATGCAGCTCCCTGTCCAAAGTCCCTTGGAGAACTCACAGCAGC 960
Qy      2953 AGCCTTAATCAAGGTTTTTAACAGCCCTTGGACACTATGAGGAGGAGGCAAGAGTACAC 3012
Db      961 AGCCTTAATCAAGGTTTTTAACAGCCCTTGGACACTATGAGGAGGAGGCAAGAGTACAC 1020
Qy      3013 CAATTTGTTTAAAGCAAGAAACCAAGTGTCTTCTCAGTGTCTTCTAGTATTTAGAAATGTTAT 3072
Db      1021 CAATTTGTTTAAAGCAAGAAACCAAGTGTCTTCTCAGTGTCTTCTAGTATTTAGAAATGTTAT 1080
Qy      3073 CATCCAGACTACTCTACCCCTGCAACATTTGAATCCCAAGAGCAAAATCCCAATTCCTCT 3132
Db      1081 CATCCAGACTACTCTACCCCTGCAACATTTGAATCCCAAGAGCAAAATCCCAATTCCTCT 1140
Qy      3133 GAGTCTGAGCTTCTGTGTAATAGGGCAGCTGTCTATGCGGTAGTAATCACATGAT 3192
Db      1141 GAGTCTGAGCTTCTGTGTAATAGGGCAGCTGTCTATGCGGTAGTAATCACATGAT 1200
Qy      3193 CTGAGGACCATTCATGGAAGTGTCTAAATAGCTAGTCTGGGAGTCTTCCATAAAGTTT 3252
Db      1201 CTGAGGACCATTCATGGAAGTGTCTAAATAGCTAGTCTGGGAGTCTTCCATAAAGTTT 1260
Qy      3253 TGCAATGGAGCAAAACAAACAGGATTTAACTAGTGTGTTTCTTCCAGCCCTCTAAAGCAT 3312
Db      1261 TGCAATGGAGCAAAACAAACAGGATTTAACTAGTGTGTTTCTTCCAGCCCTCTAAAGCAT 1320

```

QY 3313 AGGAGCTTACGCTGAGGCTTCTCTGGGCTTCTCTGTGTGTAGTATTTGTAACACTAT 3372
Db 1321 AGGGCTTAGCCTGAGGCTTCTTGGGCTTCTCTGTGTGTAGTATTTGTAACACTAT 1380
QY 3373 AGCATCTGTTAAGATCCAGTGTCCATGAGAAACCTTCCACAGCCGCTGACTGTGACTAT 3432
Db 1381 AGCATCTGTTAAGATCCAGTGTCCATGAGAAACCTTCCACAGCCGCTGACTGTGACTAT 1440
QY 3433 ATCAGTTTGGAAAGAGGCTTCTCTGCTGCTCTACAGCCGCTGAGCTGAGCTGA 3492
Db 1441 ATCAGTTTGGAAAGAGGCTTCTCTGCTGCTCTACAGCCGCTGAGCTGAGCTGA 1500
QY 3493 ATGCTTCTCTTACACCTATGTTTTAGTAGTCAAACTTCAAGAAACATCTAAACA 3552
Db 1501 ATGCTTCTCTTACACCTATGTTTTAGTAGTCAAACTTCAAGAAACATCTAAACA 1560
QY 3553 GTTCTGTTGATATGTTGTGAACTTGTATTTAGTAGTCAAACTTCAAGAAACATCTAAACA 3612
Db 1561 GTTCTGTTGATATGTTGTGAACTTGTATTTAGTAGTCAAACTTCAAGAAACATCTAAACA 1620
QY 3613 TTAACTGTTTGTGAACTGCTGATCTTCTTTCGATACCTATGATGAATAAGAA 3672
Db 1621 TTAACTGTTTGTGAACTGCTGATCTTCTTTCGATACCTATGATGAATAAGAA 1680
QY 3673 ATTAAGTGA 3682
Db 1681 ATTAAGTGA 1690

RESULT 9

US-09-738-540
; Sequence 540, Application US/09738973
; Patent No. US20020110563A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Fling, Steven P.
; APPLICANT: Mohamath, Raedoh
; APPLICANT: Algate, Paul A.
; APPLICANT: Secrist, Heather
; APPLICANT: Indritas, Carol Yoseph
; APPLICANT: Benson, Darin R.
; APPLICANT: Elliot, Mark
; APPLICANT: Mammon, Jane
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C9
; CURRENT APPLICATION NUMBER: US/09/738,973
; CURRENT FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 587
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 540
; LENGTH: 634
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(634)
; OTHER INFORMATION: n = A,T,C or G
US-09-738-973-540

Query Match 17.1%; Score 630.4; DB 9; Length 634;
Best Local Similarity 99.5%; Pred. No. 1.1e-156;
Matches 631; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2327 CCAAAAAGAGTACCAAGATTGATCTCAGCTGATGACCCCTACAGGTGCGATATGAT 2386
Db 1 CCAAAAAGAGTACCAAGATTGATCTCAGCTGATGACCCCTACAGGTGCGATATGAT 60
QY 2387 ATGAGTCTCATGGGTAAAGCAGAAAGAGTGGAAAGAAAGAACCAACCCACTCTGTCT 2446

Db 61 ATGAGTCTCATGGGTAAAGCAGAAAGAGTGGAAAGAAAGAACCAACCCACTCTGTCT 120
QY 2447 TCATATTTGATTTCAATGTTAACTCCGCTGGAATATAAGCATTCCTCTAGAGATG 2506
Db 121 TCATATTTGATTTCAATGTTAACTCCGCTGGAATATAAGCATTCCTCTAGAGATG 180
QY 2507 AGATTAAGAAAGTTTCAGATTCACAGGGGGAGAAATAGAGATTAAATCTAAAC 2566
Db 181 AGATTAAGAAAGTTTCAGATTCACAGGGGGAGAAATAGAGATTAAATCTAAAC 240
QY 2567 TGTACCTTTGGAGAGTCAATTTAAGTTAGTCCGTCTCTTTGACATCTCTGATTA 2626
Db 241 TGTACCTTTGGAGAGTCAATTTAAGTTAGTCCGTCTCTTTGACATCTCTGATTA 300
QY 2627 TTAACCCCACTCACTACCTCTGTTTCAGATGATGATGATGATGATGATGATGATGAT 2686
Db 301 TTAACCCCACTCACTACCTCTGTTTCAGATGATGATGATGATGATGATGATGATGAT 360
QY 2687 CATAGATCTCATTTGCAAGAAAGCATTTAAAGCATTCAGAAAGAAATTTAGGTTG 2746
Db 361 CATAGATCTCATTTGCAAGAAAGCATTTAAAGCATTCAGAAAGAAATTTAGGTTG 420
QY 2747 TATGACAGGCAACTGTGAGAACTGTGTGCAAAATATGATTCCTCTAGTTTTC 2806
Db 421 TATGACAGGCAACTGTGAGAACTGTGTGCAAAATATGATTCCTCTAGTTTTC 480
QY 2807 TTGTTCTCATTTGAAAGAGAAATTCACCTTTGTTTATGATTTCAAGCTTTATGATTC 2866
Db 481 TTGTTCTCATTTGAAAGAGAAATTCACCTTTGTTTATGATTTCAAGCTTTATGATTC 540
QY 2867 CATCCCATTAATAAATCTTCAAACTGCTGATGATGATGATGATGATGATGATGATGAT 2926
Db 541 CATCCCATTAATAAATCTTCAAACTGCTGATGATGATGATGATGATGATGATGATGAT 600
QY 2927 AAGTCCCTTGAGAACTCACAGAGCAGCAGGCTTA 2960
Db 601 AAGTCCCTTGAGAACTCACAGAGCAGCAGGCTTA 634

RESULT 10

US-09-854-133-540
; Sequence 540, Application US/09854133
; Publication No. US20020183499A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Raedoh
; APPLICANT: Henderson, Robert A.
; APPLICANT: Benson, Darin R.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C10
; CURRENT APPLICATION NUMBER: US/09/854,133
; CURRENT FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 735
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 540
; LENGTH: 634
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(634)
; OTHER INFORMATION: n = A,T,C or G
US-09-854-133-540

Query Match 17.1%; Score 630.4; DB 9; Length 634;
Best Local Similarity 99.5%; Pred. No. 1.1e-156;
Matches 631; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2327 CCAAAAAGAGTACCAAGATTGATCTCAGCTGATGACCCCTACAGGTGCGATATGAT 2386
Db 1 CCAAAAAGAGTACCAAGATTGATCTCAGCTGATGACCCCTACAGGTGCGATATGAT 60

Db 1 CCAAAAACAAGATGACCCAGATTTGNTTTNAGCCTGATGACCTACAGGTCGTCTATGAT 60
 2y 2387 ATGAGTCCTCATGGTAAAGCAGGAGAGGTGGGAAAGAGACACACCACTCTGTCT 2446
 Db 61 ATGAGTCCTCATGGTAAAGCAGGAGAGGTGGGAAAGAGACACACCACTCTGTCT 120
 2y 2447 TCATATTTGCAATTTTCATGTTTAACTCCGGCTGGAATAGAAAGCATTCCTTTAGAGATG 2506
 Db 121 TCATATTTGCAATTTTCATGTTTAACTCCGGCTGGAATAGAAAGCATTCCTTTAGAGATG 180
 2y 2507 AGGATAAAGAAAGTTTCAGATTTCAACAGGGGAGAGAAATGGAGATTTAACTCTAAAC 2566
 Db 181 AGGATAAAGAAAGTTTCAGATTTCAACAGGGGAGAGAAATGGAGATTTAACTCTAAAC 240
 2y 2567 TGTGACTTGGGGAGGTGAGTCATTTACAGTTAGTCTGTCTTTGACATTCCTGTGATTA 2626
 Db 241 TGTGACTTGGGGAGGTGAGTCATTTACAGTTAGTCTGTCTTTGACATTCCTGTGATTA 300
 2y 2627 TTAACCCCACTCACTACCTGTTTCAGATGATTTGGAATACCAAGATTTAAATCCTTGA 2686
 Db 301 TTAACCCCACTCACTACCTGTTTCAGATGATTTGGAATACCAAGATTTAAATCCTTGA 360
 2y 2687 CATAGATCTCATTTGAGAAAGCAGATTTAAAGACCATCAAGAGAAATTAATTTAGTTG 2746
 Db 361 CATAGATCTCATTTGAGAAAGCAGATTTAAAGACCATCAAGAGAAATTAATTTAGTTG 420
 2y 2747 TAATGCACAGCAACTGTGAGAAACTGTGTGGCCAAAATAGAAATTCCTTCTAGTTTTC 2806
 Db 421 TAATGCACAGCAACTGTGAGAAACTGTGTGGCCAAAATAGAAATTCCTTCTAGTTTTC 480
 2y 2807 TTGTTCTCATTTGAAAGAGAAATTCACATTTCTTTAGCATTTCAAGCTTTTATGTATC 2866
 Db 481 TTGTTCTCATTTGAAAGAGAAATTCACATTTCTTTAGCATTTCAAGCTTTTATGTATC 540
 2y 2867 CATCCCATCTAAAACTCTTCAAACTCCACTGTTTCTAGTCTGAAATCGAGTCCCTGTCC 2926
 Db 541 CATCCCATCTAAAACTCTTCAAACTCCACTGTTTCTAGTCTGAAATCGAGTCCCTGTCC 600
 2y 2927 AAGTGCCCTTGAGAACTCACAGCAGCGCCTTA 2960
 Db 601 AAGTGCCCTTGAGAACTCACAGCAGCGCCTTA 634

RESULT 11
 US-10-144-649A-540
 ; Sequence 540, Application US/10144649A
 ; Publication No. US20030118599A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lodes, Michael J.
 ; APPLICANT: Wang, Tongtong
 ; APPLICANT: Fan, Liqun
 ; APPLICANT: Algate, Paul A.
 ; APPLICANT: McNeill, Patricia D.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
 ; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
 ; FILE REFERENCE: 210121.475C11
 ; CURRENT APPLICATION NUMBER: US/10/144,649A
 ; CURRENT FILING DATE: 2002-08-21
 ; NUMBER OF SEQ ID NOS: 749
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 540
 ; LENGTH: 634
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: 25..29
 ; OTHER INFORMATION: n = A,T,C or G
 US-10-144-649A-540
 Query Match 17.1%; Score 630.4; DB 14; Length 634;
 Best Local Similarity 99.5%; Pred. No. 1.1e-156;
 Matches 631; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2327 CCAAAAACAAGATGACCCAGATTTGATCTCAGCCTGATGACCTACAGTCTGTCTATGAT 2386
 Db 1 CCAAAAACAAGATGACCCAGATTTGATCTCAGCCTGATGACCTACAGTCTGTCTATGAT 60
 QY 2387 ATGAGTCCTCATGGTAAAGCAGGAGAGGTGGGAAAGAGACACACCACTCTGTCT 2446
 Db 61 ATGAGTCCTCATGGTAAAGCAGGAGAGGTGGGAAAGAGACACACCACTCTGTCT 120
 QY 2447 TCATATTTGCAATTTTCATGTTTAACTCCGGCTGGAATAGAAAGCATTCCTTTAGAGATG 2506
 Db 121 TCATATTTGCAATTTTCATGTTTAACTCCGGCTGGAATAGAAAGCATTCCTTTAGAGATG 180
 QY 2507 AGGATAAAGAAAGTTTCAGATTTCAACAGGGGAGAGAAATGGAGATTTAACTCTAAAC 2566
 Db 181 AGGATAAAGAAAGTTTCAGATTTCAACAGGGGAGAGAAATGGAGATTTAACTCTAAAC 240
 QY 2567 TGTGACTTGGGGAGGTGAGTCATTTACAGTTAGTCTGTCTTTGACATTCCTGTGATTA 2626
 Db 241 TGTGACTTGGGGAGGTGAGTCATTTACAGTTAGTCTGTCTTTGACATTCCTGTGATTA 300
 QY 2627 TTAACCCCACTCACTACCTGTTTCAGATGATTTGGAATACCAAGATTTAAATCCTTGA 2686
 Db 301 TTAACCCCACTCACTACCTGTTTCAGATGATTTGGAATACCAAGATTTAAATCCTTGA 360
 QY 2687 CATAGATCTCATTTGAGAAAGCAGATTTAAAGACCATCAAGAGAAATTAATTTAGTTG 2746
 Db 361 CATAGATCTCATTTGAGAAAGCAGATTTAAAGACCATCAAGAGAAATTAATTTAGTTG 420
 QY 2747 TAATGCACAGCAACTGTGAGAAACTGTGTGGCCAAAATAGAAATTCCTTCTAGTTTTC 2806
 Db 421 TAATGCACAGCAACTGTGAGAAACTGTGTGGCCAAAATAGAAATTCCTTCTAGTTTTC 480
 QY 2807 TTGTTCTCATTTGAAAGAGAAATTCACATTTCTTTAGCATTTCAAGCTTTTATGTATC 2866
 Db 481 TTGTTCTCATTTGAAAGAGAAATTCACATTTCTTTAGCATTTCAAGCTTTTATGTATC 540
 QY 2867 CATCCCATCTAAAACTCTTCAAACTCCACTGTTTCTAGTCTGAAATCGAGTCCCTGTCC 2926
 Db 541 CATCCCATCTAAAACTCTTCAAACTCCACTGTTTCTAGTCTGAAATCGAGTCCCTGTCC 600
 QY 2927 AAGTGCCCTTGAGAACTCACAGCAGCGCCTTA 2960
 Db 601 AAGTGCCCTTGAGAACTCACAGCAGCGCCTTA 634

RESULT 12
 US-09-735-705-209
 ; Sequence 209, Application US/09735705
 ; Patent No. US20020052329A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Tongtong
 ; APPLICANT: Fan, Liqun
 ; APPLICANT: Kalos, Michael D.
 ; APPLICANT: Bangur, Chaitanya S.
 ; APPLICANT: Hosken, Nancy R.
 ; APPLICANT: Fanger, Gary R.
 ; APPLICANT: Li, Samuel X.
 ; APPLICANT: Wang, Aijun
 ; APPLICANT: Skeiky, Yasir A.W.
 ; APPLICANT: Henderson, Robert A.
 ; APPLICANT: McNeill, Patricia D.
 ; APPLICANT: Fanger, Neil
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
 ; FILE REFERENCE: 210121.455C14
 ; CURRENT APPLICATION NUMBER: US/09/735,705
 ; CURRENT FILING DATE: 2000-12-12
 ; NUMBER OF SEQ ID NOS: 419
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 209
 ; LENGTH: 621
 ; TYPE: DNA

ORGANISM: Homo sapiens
US-09-735-705-209

Query Match 16.8%; Score 619.4; DB 9; Length 621;
Best Local Similarity 99.8%; Pred. No. 9.1e-154;
Matches 620; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3055 CATTAGAACATGGTTATCATCCAGACTCTACCCCTGCAACATTGAACTGCCAAG 3114
DB 1 CATTAGAACATGGTTATCATCCAGACTCTACCCCTGCAACATTGAACTGCCAAG 60
QY 3115 CAATCCACATTCCTCTTGAATCTGACGCTCTGTGTAAATAGGGACGCTGTCTAT 3174
DB 61 CAATCCACATTCCTCTTGAATCTGACGCTCTGTGTAAATAGGGACGCTGTCTAT 120
QY 3175 GCCGTAAATCAGATGATCTGAGACCAATTCATGGAAGCTGTAAATAGCTGTGG 3234
DB 121 GCCGTAAATCAGATGATCTGAGACCAATTCATGGAAGCTGTAAATAGCTGTGG 180
QY 3235 GAGTCTTCATTAAGTTTGGATGAGCAAAACAAGATTAACCTAGCTTGTCT 3294
DB 181 GAGTCTTCATTAAGTTTGGATGAGCAAAACAAGATTAACCTAGCTTGTCT 240
QY 3295 TCAGCCCTCTAAAGCATAGGGCTTAGCCTGACAGGCTTCTTGGGCTTCTGTGTG 3354
DB 241 TCAGCCCTCTAAAGCATAGGGCTTAGCCTGACAGGCTTCTTGGGCTTCTGTGTG 300
QY 3355 TAGTTTGTAAACATATAGCATCTGTTAAGATCCAGTGTCAATGAAACCTTCCACAT 3414
DB 301 TAGTTTGTAAACATATAGCATCTGTTAAGATCCAGTGTCAATGAAACCTTCCACAT 360
QY 3415 GCCGTGACTGGAATATAGTTTGGAAAGCAGGTTCTCTGCTGCTAAACAAG 3474
DB 361 GCCGTGACTGGAATATAGTTTGGAAAGCAGGTTCTCTGCTGCTAAACAAG 420
QY 3475 CCAGTGGACAGCTGATGATCTTCTCTTACACATGTTTAAAGTAAACCTTCC 3534
DB 421 CCAGTGGACAGCTGATGATCTTCTCTTACACATGTTTAAAGTAAACCTTCC 480
QY 3535 AAGAAACATCTAAACAAGTTCTGTGATATGTGTGAACTGTGATTTA 3594
DB 481 AAGAAACATCTAAACAAGTTCTGTGATATGTGTGAACTGTGATTTA 540
QY 3595 GTAGGCTCTATATGATTAACCTGTTTGTAACTCCGATCTTCTTCCGATA 3654
DB 541 GTAGGCTCTATATGATTAACCTGTTTGTAACTCCGATCTTCTTCCGATA 600
QY 3655 CTATTGATGAATAAAGAAATT 3675
DB 601 CTATTGATGAATAAAGAAATT 621

RESULT 13
US-09-850-716A-209

Sequence 209, Application US/09850716A
Patent No. US20020115139A1
GENERAL INFORMATION:
APPLICANT: Kalo, Michael D.
APPLICANT: McNeill, Patricia D.
APPLICANT: Retter, Marc W.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 210121.455C15
CURRENT APPLICATION NUMBER: US/09/850.716A
CURRENT FILING DATE: 2001-05-07
NUMBER OF SEQ ID NOS: 440
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 209
LENGTH: 621
TYPE: DNA
ORGANISM: Homo sapiens
US-09-850-716A-209

Query Match 16.8%; Score 619.4; DB 9; Length 621;
Best Local Similarity 99.8%; Pred. No. 9.1e-154;
Matches 620; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3055 CATTAGAACATGGTTATCATCCAGACTCTACCCCTGCAACATTGAACTGCCAAG 3114
DB 1 CATTAGAACATGGTTATCATCCAGACTCTACCCCTGCAACATTGAACTGCCAAG 60
QY 3115 CAATCCACATTCCTCTTGAATCTGACGCTCTGTGTAAATAGGGACGCTGTCTAT 3174
DB 61 CAATCCACATTCCTCTTGAATCTGACGCTCTGTGTAAATAGGGACGCTGTCTAT 120
QY 3175 GCCGTAAATCAGATGATCTGAGACCAATTCATGGAAGCTGTAAATAGCTGTGG 3234
DB 121 GCCGTAAATCAGATGATCTGAGACCAATTCATGGAAGCTGTAAATAGCTGTGG 180
QY 3235 GAGTCTTCATTAAGTTTGGATGAGCAAAACAAGATTAACCTAGCTTGTCT 3294
DB 181 GAGTCTTCATTAAGTTTGGATGAGCAAAACAAGATTAACCTAGCTTGTCT 240
QY 3295 TCAGCCCTCTAAAGCATAGGGCTTAGCCTGACAGGCTTCTTGGGCTTCTGTGTG 3354
DB 241 TCAGCCCTCTAAAGCATAGGGCTTAGCCTGACAGGCTTCTTGGGCTTCTGTGTG 300
QY 3355 TAGTTTGTAAACATATAGCATCTGTTAAGATCCAGTGTCAATGAAACCTTCCACAT 3414
DB 301 TAGTTTGTAAACATATAGCATCTGTTAAGATCCAGTGTCAATGAAACCTTCCACAT 360
QY 3415 GCCGTGACTGGAATATAGTTTGGAAAGCAGGTTCTCTGCTGCTAAACAAG 3474
DB 361 GCCGTGACTGGAATATAGTTTGGAAAGCAGGTTCTCTGCTGCTAAACAAG 420
QY 3475 CCAGTGGACAGCTGATGATCTTCTCTTACACATGTTTAAAGTAAACCTTCC 3534
DB 421 CCAGTGGACAGCTGATGATCTTCTCTTACACATGTTTAAAGTAAACCTTCC 480
QY 3535 AAGAAACATCTAAACAAGTTCTGTGATATGTGTGAACTGTGATTTA 3594
DB 481 AAGAAACATCTAAACAAGTTCTGTGATATGTGTGAACTGTGATTTA 540
QY 3595 GTAGGCTCTATATGATTAACCTGTTTGTAACTCCGATCTTCTTCCGATA 3654
DB 541 GTAGGCTCTATATGATTAACCTGTTTGTAACTCCGATCTTCTTCCGATA 600
QY 3655 CTATTGATGAATAAAGAAATT 3675
DB 601 CTATTGATGAATAAAGAAATT 621

RESULT 14
US-09-897-778-209

Sequence 209, Application US/09897778
Patent No. US20020147143A1
GENERAL INFORMATION:
APPLICANT: Wang, Tonglong
APPLICANT: Wainerakis, Margarita
APPLICANT: Fanger, Gary R.
APPLICANT: Veddyk, Thomas S.
APPLICANT: Carter, Darrick
APPLICANT: Matanabe, Yoshihiro
APPLICANT: Henderson, Robert A.
APPLICANT: Beckham, David M.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 210121.455C16
CURRENT APPLICATION NUMBER: US/09/897.778
CURRENT FILING DATE: 2001-06-28
NUMBER OF SEQ ID NOS: 467
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 209
LENGTH: 621
TYPE: DNA

ORGANISM: Homo sapiens
JS-09-897-778-209

Query Match 16.8%; Score 619.4; DB 9; Length 621;
Best Local Similarity 99.8%; Pred. No. 9.1e-154;
Matches 620; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

2y 3055 CATTAGAACATGGTTATCATCCAGACTACTACCTGCAACATGAACCTCCAGAG 3114
2b 1 CATTAGAACATGGTTATCATCCAGACTACTACCTGCAACATGAACCTCCAGAG 60

2y 3115 CAAATCCACATTCCTCTGAGTCTGAGCTTCTGTGTAATAGGCGAGCTGCTCTAT 3174
2b 61 CAAATCCACATTCCTCTGAGTCTGAGCTTCTGTGTAATAGGCGAGCTGCTCTAT 120

2y 3175 GCCGTAGAACATCATGATCTGAGACCAATCATGGAAGCTGCTAAATAGCTAGTCTGG 3234
2b 121 GCCGTAGAACATCATGATCTGAGACCAATCATGGAAGCTGCTAAATAGCTAGTCTGG 180

2y 3235 GAGTCTTCCATAAAGTTTTCATGAGCAAAACAAACAGGATTAACCTAGGTTGGTTCCT 3294
2b 181 GAGTCTTCCATAAAGTTTTCATGAGCAAAACAAACAGGATTAACCTAGGTTGGTTCCT 240

2y 3295 TCAGCCCTCTAAAGCATAGGGCTTAGCCTGAGGCTTCCCTGAGGCTTCTCTGTGTG 3354
2b 241 TCAGCCCTCTAAAGCATAGGGCTTAGCCTGAGGCTTCCCTGAGGCTTCTCTGTGTG 300

2y 3355 TAGTTTTGTAACACTATAGCATCTGTTAAGATCCAGTGTCCATGGAACCTTCCACAT 3414
2b 301 TAGTTTTGTAACACTATAGCATCTGTTAAGATCCAGTGTCCATGGAACCTTCCACAT 360

2y 3415 GCCGTGACTGGAATATATAGTTTGGAAAGCAGGGTTCCTGTGCTGCTAACAGC 3474
2b 361 GCCGTGACTGGAATATATAGTTTGGAAAGCAGGGTTCCTGTGCTGCTAACAGC 420

2y 3475 CCAGTGGACAGTCTGATGCTTCCCTTTTACACCTATGTTTAAAGTAGTCAAACTTC 3534
2b 421 CCAGTGGACAGTCTGATGCTTCCCTTTTACACCTATGTTTAAAGTAGTCAAACTTC 480

2y 3535 AAGAAACAATCTAAACAAGTTTCTGTCATATGTTTGTGAACCTTGTATTGATTTA 3594
2b 481 AAGAAACAATCTAAACAAGTTTCTGTCATATGTTTGTGAACCTTGTATTGATTTA 540

2y 3595 GTAGGCTTCTATATGCAATTAACCTGTTTGTAACTCCCTGATTCCTCTTCGGATA 3654
2b 541 GTAGGCTTCTATATGCAATTAACCTGTTTGTAACTCCCTGATTCCTCTTCGGATA 600

2y 3655 CTATTGATGAATAAAGAAATT 3675
2b 601 CTATTGATGAATAAAGAAATT 621

RESULT 15
US-09-466-396A-209
Sequence 209, Application US/09466396A
Publication No. US20030119763A1
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY AND
FILE OF INVENTION: DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.455C4
CURRENT APPLICATION NUMBER: US/09/466.396A
CURRENT FILING DATE: 1999-12-17
NUMBER OF SEQ ID NOS: 224
SOFTWARE: Fast-Seq for Windows Version 3.0
SEQ ID NO 209
LENGTH: 621
TYPE: DNA
ORGANISM: Homo sapiens
US-09-466-396A-209

Query Match 16.8%; Score 619.4; DB 10; Length 621;
Best Local Similarity 99.8%; Pred. No. 9.1e-154;

Matches 620; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3055 CATTAGAACATGGTTATCATCCAGACTACTACCTGCAACATGAACCTCCAGAG 3114
DB 1 CATTAGAACATGGTTATCATCCAGACTACTACCTGCAACATGAACCTCCAGAG 60

QY 3115 CAAATCCACATTCCTCTGAGTCTGAGCTTCTGTGTAATAGGCGAGCTGCTCTAT 3174
DB 61 CAAATCCACATTCCTCTGAGTCTGAGCTTCTGTGTAATAGGCGAGCTGCTCTAT 120

QY 3175 GCCGTAGAACATCATGATCTGAGACCAATCATGGAAGCTGCTAAATAGCTAGTCTGG 3234
DB 121 GCCGTAGAACATCATGATCTGAGACCAATCATGGAAGCTGCTAAATAGCTAGTCTGG 180

QY 3235 GAGTCTTCCATAAAGTTTTCATGAGCAAAACAAACAGGATTAACCTAGGTTGGTTCCT 3294
DB 181 GAGTCTTCCATAAAGTTTTCATGAGCAAAACAAACAGGATTAACCTAGGTTGGTTCCT 240

QY 3295 TCAGCCCTCTAAAGCATAGGGCTTAGCCTGAGGCTTCCCTGAGGCTTCTCTGTGTG 3354
DB 241 TCAGCCCTCTAAAGCATAGGGCTTAGCCTGAGGCTTCCCTGAGGCTTCTCTGTGTG 300

QY 3355 TAGTTTTGTAACACTATAGCATCTGTTAAGATCCAGTGTCCATGGAACCTTCCACAT 3414
DB 301 TAGTTTTGTAACACTATAGCATCTGTTAAGATCCAGTGTCCATGGAACCTTCCACAT 360

QY 3415 GCCGTGACTGGAATATATAGTTTGGAAAGCAGGGTTCCTGTGCTGCTAACAGC 3474
DB 361 GCCGTGACTGGAATATATAGTTTGGAAAGCAGGGTTCCTGTGCTGCTAACAGC 420

QY 3475 CCAGTGGACAGTCTGATGCTTCCCTTTTACACCTATGTTTAAAGTAGTCAAACTTC 3534
DB 421 CCAGTGGACAGTCTGATGCTTCCCTTTTACACCTATGTTTAAAGTAGTCAAACTTC 480

QY 3535 AAGAAACAATCTAAACAAGTTTCTGTCATATGTTTGTGAACCTTGTATTGATTTA 3594
DB 481 AAGAAACAATCTAAACAAGTTTCTGTCATATGTTTGTGAACCTTGTATTGATTTA 540

QY 3595 GTAGGCTTCTATATGCAATTAACCTGTTTGTAACTCCCTGATTCCTCTTCGGATA 3654
DB 541 GTAGGCTTCTATATGCAATTAACCTGTTTGTAACTCCCTGATTCCTCTTCGGATA 600

QY 3655 CTATTGATGAATAAAGAAATT 3675
DB 601 CTATTGATGAATAAAGAAATT 621

Search completed: March 12, 2004, 06:13:28
Job time : 865.651 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

DM nucleic - nucleic search, using sw model

Run on: March 10, 2004, 09:18:38 ; Search time 6114.16 seconds
(without alignments)
18036.975 Million cell updates/sec

Title: US-10-084-817-20

Perfect score: 3693

Sequence: 1 cgcacggcgactccaccct.....ttaaagtgaataaaaaaa 3693

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- EST:
- 1: em_estba:**
 - 2: em_esthum:**
 - 3: em_estin:**
 - 4: em_estmu:**
 - 5: em_estov:**
 - 6: em_estpl:**
 - 7: em_estro:**
 - 8: em_hcc:**
 - 9: gb_est1:**
 - 10: gb_est2:**
 - 11: gb_hcc:**
 - 12: gb_est3:**
 - 13: gb_est4:**
 - 14: gb_est5:**
 - 15: em_estfun:**
 - 16: em_estom:**
 - 17: em_gss_hum:**
 - 18: em_gss_inv:**
 - 19: em_gss_pln:**
 - 20: em_gss_vrt:**
 - 21: em_gss_fun:**
 - 22: em_gss_mam:**
 - 23: em_gss_mus:**
 - 24: em_gss_pro:**
 - 25: em_gss_rod:**
 - 26: em_gss_phg:**
 - 27: em_gss_vrl:**
 - 28: gb_gss1:**
 - 29: gb_gss2:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1050	28.4	3610	11 AK049440	AK049440 Mus muscu
2	1044.6	28.3	1110	9 AL548992	AL548992 AL548992
3	1007.6	27.3	1109	9 AL573353	AL573353 AL573353
4	991.2	26.8	1090	12 BM553173	BM553173 AGENCOURT

5	988.6	26.8	1094	12	BM463192	BM463192 AGENCOURT
6	960.4	26.0	1677	11	BC006185	BC006185 Mus muscu
7	913.6	24.7	999	9	AL570818	AL570818 AL570818
8	905.6	24.5	1048	13	EX338760	EX338760 EX338760
9	905.6	24.5	1201	13	EX405684	EX405684 EX405684
10	879.6	23.8	995	9	AL544661	AL544661 AL544661
11	875.6	23.7	987	13	EX395667	EX395667 EX395667
12	869.6	23.5	1201	13	EX352901	EX352901 EX352901
13	845.8	22.9	916	13	EX0671550	EX0671550 AGENCOURT
14	844	22.9	879	9	AB045227	AB045227 AB045227
15	844	22.9	927	13	EX018886	EX018886 AGENCOURT
16	842.8	22.8	906	13	BU146207	BU146207 AGENCOURT
17	840.8	22.8	917	13	BQ718490	BQ718490 AGENCOURT
18	838.8	22.7	891	14	CF242911	CF242911 AGENCOURT
19	835	22.6	1201	13	EX405683	EX405683 EX405683
20	825.4	22.4	1068	12	BM912228	BM912228 AGENCOURT
21	821.4	22.2	878	14	CD518823	CD518823 AGENCOURT
22	821.4	22.2	911	13	BQ883171	BQ883171 AGENCOURT
23	814.8	22.1	886	13	BQ932018	BQ932018 AGENCOURT
24	807.4	21.9	1142	12	BM916732	BM916732 AGENCOURT
25	801.2	21.7	853	13	BU517310	BU517310 AGENCOURT
26	797.2	21.6	829	9	AU140319	AU140319 AU140319
27	795.2	21.5	933	13	BQ897777	BQ897777 AGENCOURT
28	791.2	21.4	809	9	AU130022	AU130022 AU130022
29	791.2	21.4	892	13	BQ707255	BQ707255 AGENCOURT
30	789.8	21.4	831	9	AU131134	AU131134 AU131134
31	786.4	21.3	833	9	AU132046	AU132046 AU132046
32	784.4	21.2	904	14	CD244996	CD244996 AGENCOURT
33	783.2	21.2	840	12	BG570724	BG570724 602591838
34	781.4	21.2	872	14	CD515270	CD515270 AGENCOURT
35	780.6	21.1	929	13	BQ722506	BQ722506 AGENCOURT
36	778	21.1	799	14	CK000721	CK000721 AGENCOURT
37	773.6	20.9	875	14	CD388013	CD388013 AGENCOURT
38	772.6	20.9	964	13	BQ707449	BQ707449 AGENCOURT
39	771.4	20.9	1030	13	BQ952029	BQ952029 AGENCOURT
40	771.2	20.9	910	13	BQ719949	BQ719949 AGENCOURT
41	769.8	20.8	876	14	CD654748	CD654748 AGENCOURT
42	767	20.8	885	9	AL710894	AL710894 DKFZp686G
43	766.6	20.8	801	14	CK000499	CK000499 AGENCOURT
44	766.4	20.8	1101	9	AL574665	AL574665 AL574665
45	756.2	20.5	773	14	CA414235	CA414235 UI-H-E20-

ALIGNMENTS

RESULT 1	AK049440	3610 bp	mRNA	linear	HTC 20-SEP-2003
LOCUS	Mus musculus	7 days embryo whole body cDNA, RIKEN full-length			
DEFINITION	enriched library, clone:CA3001K03	product:follicle-statin-like, full insert sequence.			
ACCESSION	AK049440				
VERSION	AK049440.1	GI:26340175			
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus	(house mouse)			
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.				
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.				
TITLE	High-efficiency full-length cDNA cloning				
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)				
MEDLINE	99279253				
PUBMED	10349636				
REFERENCE	1				
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.				
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)				
MEDLINE	20499374				
PUBMED	11042159				


```
863 GTGCCACTGGTCTGCACAGCAATACCTGTGATGGAAGAATCAGAAGGGGGTCCAGA 922
912 CCAGACAGAGGAGGAGATGACAGATATGTCTCAGAGCTCCAAAAGCATCAGGAACAG 971
923 CCACACAGAGGAGGAGAGAGGATATGTCCAGGAATCCAGAAACACAGGGGACAG 982
972 CTGAAAGACCAAGAGAGTGAGCACCAGAGATCTAATGAG-GAGGACACAGACAGTGT 1030
983 CAGAAAGACCAAGAGAGTGAACACCAAGAGAGATCTAAGAGAGGACACAGACCGTGT 1042
1031 CTGATGCCAGCATCTTCTCCATCTCAGCGCTGAGTTCAGTATACACAGAGTCTGCTAC 1090
1043 CCGAGGCCAGCGCCT---CCTCTCAGCGCTGAGCCAGTACACAC-AGAGTCTGCAGC 1098
1091 AGTCGCCAAATCACCAGTATTTCTTATATAGCAATGAGTTTATTTTATTTTATTT 1150
1099 AATCACCANAATCACTAGTATTTCTTGTATGGCAGCAATCTTATTTG---TTGTTT 1154
1151 TCGAATAAGGATATGAGTGTGCTGCTAGGAGGAGGAGGACAGGCTTCATTTCTA 1210
1155 TGCATAAAGGAATGAGGTGGCCAGCTAGCGAGGAGGAGGCCAACCTTCACCTGTA 1214
1211 GGAAGTCTTAAAGAGAACTGTAAATGGTCTCTGGGGCTGAGGCTAGTAAGGAACTG 1270
1215 GGAATGCTTAAAGAGAAC--TAAGGACACCTTGGAGAGAGGCAACTAAGGAACAG 1272
1271 CATCAGATTAAAGAGAGAACAGACCCAAATCTGAAC---CTCTTTGAGTTACTGCA 1326
1273 CATC-GGGTTGGCAGAGGAGCAGAGGCGAGTTTGAATGAAGCCCTTCGGGGTCACAG 1331
1327 TCTGTACAGAGGCTCAGGAGGTGCACACGATGCCAGAGAGAACTTAGCAGGCTGTC 1386
1332 CTGCGAGGAGATACAGGAAGCAT-----AGGAACACTTGAAGTCCAGTCCCTG 1381
1387 GAGAGGAGGTTTGGAGAGTCCACGAGAGGAAACGCTCTCTGCTTCCAGCCTCTTTCCA 1446
1382 CTGAGGAGAGTGGGGGAGCTTTGTAGGAGGAAACCTCTGCTGTTT---GACCCCTGTCA 1438
1447 TTGCGGTGACATGACAGACCTCAGCATCCAGCATCTCTGGTCCCAATAC-TGCCT 1505
1439 CCACTGTCACTGACAGACCTGACGCAAGTCTGCTTCTCTTTGGTCCCAACATCAC 1498
1506 CTGATACATAGCCATCTGCTAGTTAAACCAAGTCCCTCAGACTTGGATGAGTTCT 1565
1499 CTGAACACACAGCGGCCCAACTAGTTA--CCTGTGCTCCTCAGCCTTGCATGAGTTCC 1556
1566 GGGAGGTACACCAATGATGACATCTGTATCTTGTAGCCCTTAGCGACCTAAC 1625
1557 TGGAGGAGGTTTAAATGATGACAGACCTTATGATCTCAAGCGCATGGAGACCTAAC 1616
1626 CAAATTTTAAAAATACTTTT-----ACCAAGGTGCTATTTCT 1664
1617 CAAATTTTAAAAATACATTTTCTTTTCTTTTGTGTTAACCAAGGTGCTATTTCT 1676
1665 CTGTAAACACTTTTTCGCAAGTGAATTTCTTCAATTTATATATATATATAT 1724
1677 CTGTAAAGAGCTTTT---CCAGCTGACTTCATTCCTCAGTATATACCGTTATATAT 1733
1725 TGT--TTTTTAATATTTTATTTTCTGACTAGGTATTAAGCTTTTGTAAATTA-TTTTCA 1781
1734 TGTGTTTTTAAATATTTTCAATTTTGTAGTATATTAAGCTTTTGTAAATTTTCA 1793
1782 GTAGTCCACACTTCTAGGTGA-----AGGAGTTGGGGTTCTTCCTGGTGACAG 1834
1794 TTAGTCTACTATTTTCAGAGTGAAGGTGAAGGGGTTTGGGCATTTTCCAGGTACAGG 1853
1835 GGCTGAAATAACCCAGA-TGCTCCACCTGCCACATACCTAGTA-----TGCAGCCCATAG 1888
1854 GAATCTGTAAACACACAGCCCATACCTCTGCATATTTAGACCGGTTGCAGTTCGGAG 1913
1889 TTGGCC-----CCCTAGTCTCAGCAGTCCA---CTATCTGCAGAGGACAGAGGTGC 1940
```

```
1914 CTGACCCCAACCCAGAGCTTTCTAGAAAATCAGCTCCATGCCAGAAACCAAGAGGCC 1973
1941 CTTAGACTTAAGCCAGGGAGAGCATCTTTCAAAAAAATCTTTCAAGATCCCAACA-TT 1999
1974 CTCAGACAGAAGCCACAGGACCAAGCATCTTCATAGACAGCTGTTGAGATCCCAACAGTT 2033
2000 AATTGTTTTTATTTATTTATCTCGAAGTT 2027
2034 AATTGCTTTTGTCTTCTTGAAGAATT 2061

RESULT 2
AL548992
LOCUS
DEFINITION
AL548992 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1042YC07 5-PRIME, mRNA sequence.
ACCESSION
AL548992
VERSION
AL548992.2 GI:31270813
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1110)
AUTHORS
Li, W.B., Gruber, C., Jeesee, J. and Polayes, D.
TITLE
Full-length cDNA libraries and normalization
JOURNAL
Unpublished (2001)
COMMENT
On Feb 15, 2001 this sequence version replaced gi:12884545.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seque@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 1910.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0D1042AB04QPI&cluster=1910.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0D1042AB04QPI.
FEATURES
Location/Qualifiers
1..1110
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1042YC07"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
Query Match 28.3%; Score 1044.6; DB 9; Length 1110;
Best Local Similarity 99.3%; Pred. No. 1.6e-199;
Matches 1044; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
QY 2 GATCGGCGACTCCACCTCGCTTACAGCTCGCTCGCGCGTCTCGCGCGCCCA 61
DB 60 GATCGGCGACTCCACCTCGCTTACAGCTCGCTCGCGCGTCTCGCGCGCCCA 119
QY 62 GGAGACCTGGACACAGACAGTGTGAAACGCTCGGCTCGGCTCGGCTCGGCTCGG 121
DB 120 GGAGACCTGGACACAGACAGTGTGAAACGCTCGGCTCGGCTCGGCTCGGCT 179
QY 122 CGGCTCGGCTCGGCTCGGCGGAGAGCTAAGAGAGCAATCCAAAGATCTGTGCAAT 181
DB 180 CGGCTCGGCTCGGCTCGGCGGAGAGCTAAGAGAGCAATCCAAAGATCTGTGCAAT 239
QY 182 GTGTTTTGTGAGCGCGCGGGAATGTGCAGTACACAGAAAGGGGAACCCACTGTCTC 241
DB 240 GTGTTTTGTGAGCGCGCGGGAATGTGCAGTACACAGAAAGGGGAACCCACTGTCTC 299
```



```

492 GTCTGGGGAGTCTCCATAAGTTTTCATGGGAGCAAAACAGGATTAAACATGGTTT 433
3288 GGTTCCTTCAGCCCTCTAAAGCATAGGGCTTAGCCTGAGGGCTTCCTTGGGCTTCTCT 3347
432 GGTTCCTTCAGCCCTCTAAAGCATAGGGCTTAGCCTGAGGGCTTCCTTGGGCTTCTCT 373
3348 GTGTGTAGTATTTTGAACACATATAGCTCTGTTAAGATCCAGTGTCCATGGAACCTT 3407
372 GTGTGTAGTATTTTGAACACATATAGCTCTGTTAAGATCCAGTGTCCATGGAACAT 313
3408 CCCACATGCCGTGACTCTGACATATATCAGTTTGGAAAGCAGGGTTCCTCTGCCCTGCT 3467
312 CCCACATGCCGTGACTCTGACATATATCAGTTTGGAAAGCAGGGTTCCTCTGCCCTGCT 253
3468 AACAGCCACGTGGACAGTCTGAATGTCTTCCCTTTACACCTATGTTTTHAAGTAGTC 3527
252 AACAGCCACGTGGACAGTCTGAATGTCTTCCCTTTACACCTATGTTTTHAAGTAGTC 193
3528 AAACCTTCAAGAAACAATCTAAACAAGTTTCTGTGTGATATGTTTGTGAACCTTGATTT 3587
192 AAACCTTCAAGAAACAATCTAAACAAGTTTCTGTGTGATATGTTTGTGAACCTTGATTT 133
3588 GTATTAGTAGGCTCTATATGATTAATTAACCTTGTGTTTGTAACTCTGATTCCTTT 3647
132 GTATTAGTAGGCTCTATATGATTAATTAACCTTGTGTTTGTAACTCTGATTCCTTT 73
3648 TCGGATACTATTGATGAATAAAGAAATTA 3677
72 TCGGATACTATTGATGAATAAAGAAATTA 43

```

```

RESULT 4
BM553173 1090 bp mRNA linear EST 20-FEB-2002
LOCUS AGENCOURT 6572495 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5467082
5' mRNA sequence.
BM553173
VERSION BM553173.1 GI:18791672
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1090)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM1968 row: i column: 03
High quality sequence stop: 690.
Location/Qualifiers
1..1090
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5467082"
/tissue_type="amelanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_41"
/notes="Organ: Skin; Vector: pOTB7; Site: 1: XhoI; Site 2:
EcoRI; CDNA made by oligo-dr priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCACGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,

```

FEATURES

source

```

ORIGIN
Query Match 26.8%; Score 991.2; DB 12; Length 1090;
Best Local Similarity 97.2%; Pred. No. 8.7e-189;
Matches 1049; Conservative 0; Mismatches 26; Indels 4; Gaps 4;

QY 537 ATTTTAAAGAACTTTGATATGTTGATTTCTCCCTGGAGTCCAGTGAATTCCTGAAGTTTG 596
Db 1 ATTTTAAAGAACTTTGATATGTTGATTTCTCCCTGGAGTCCAGTGAATTCCTGAAGTTTG 60
QY 597 TGGACACAATGAAGAACTGCCATCAATATTACAAGTATCCAGACGAGAGAACCAAGT 656
Db 61 TGGACACAATGAAGAACTGCCATCAATATTACAAGTATCCAGACGAGAGAACCAAGT 120
QY 657 TGCTTAGGGAGTCTGTGTTGATGCTCTCAATGAACCTGTCTGATGAAATGCTGATTGGA 716
Db 121 TGCTTAGGGAGTCTGTGTTGATGCTCTCAATGAACCTGTCTGATGAAATGCTGATTGGA 180
QY 717 AACTCAGCTTCCAAAGAGTTTCTCAAGTGCCTCAACCCATCTTCAACCCCTCCTGAGAGA 776
Db 181 AACTCAGCTTCCAAAGAGTTTCTCAAGTGCCTCAACCCATCTTCAACCCCTCCTGAGAGA 240
QY 777 AGTGTGCCCTCGAGAGTGAAGAACTGATGAGATGAGCTGAGACCGAGGTGACTGTAAAC 836
Db 241 AGTGTGCCCTCGAGAGTGAAGAACTGATGAGATGAGCTGAGACCGAGGTGACTGTAAAC 300
QY 837 GCTGTGCTGTGCTGTGGAAATTTGGGTCTGTACAGCATGACCTGTGACGGAAGAAATC 896
Db 301 GCTGTGCTGTGCTGTGGAAATTTGGGTCTGTACAGCATGACCTGTGACGGAAGAAATC 360
QY 897 AGAAGGGGGCCAGACCCAGACAGAGGAGGAGATGACCAAGATATGTCCAGAGGTCCAAA 956
Db 361 AGAAGGGGGCCAGACCCAGACAGAGGAGGAGATGACCAAGATATGTCCAGAGGTCCAAA 420
QY 957 AGCATCAGAAACAGCTGAAAAGCCAGAGAGTGAACCCACCAAGAGATCTAATGAGAG 1016
Db 421 AGCATCAGAAACAGCTGAAAAGCCAGAGAGTGAACCCACCAAGAGATCTAATGAGAG 480
QY 1017 GCACAGACAGCTGTGATCCCGATCTTCTCCACTTTCAGCGCTGAGTTTCAGTATACA 1076
Db 481 GCACAGACAGCTGTGATCCCGATCTTCTCCACTTTCAGCGCTGAGTTTCAGTATACA 540
QY 1077 CAAAGTCTGTACAGTGCCTCAATCACCAGTATTTGCTTATATAGCAATGAGTTTATT 1136
Db 541 CAAAGTCTGTACAGTGCCTCAATCACCAGTATTTGCTTATATAGCAATGAGTTTATT 600
QY 1137 TTGTTTATTGTTTGGCAATAAAGGATATGAAGTGGCTGGCTAGGAAGGGAAGGGCCAC 1196
Db 601 TTGTTTATTGTTTGGCAATAAAGGATATGAAGTGGCTGGCTAGGAAGGGAAGGGCCAC 660
QY 1197 AGCCTTCATTTCTAGGAGTGTCTTAAAGAAACCTGTAAATGGTCTCTGCGGCTGGAGGC 1256
Db 661 AGCCTTCATTTCTAGGAGTGTCTTAAAGAAACCTGTAAATGGTCTCTGCGGCTGGAGGC 720
QY 1257 TAGTAAGGAAACTGCATCAGATGAAAGAGGAAACAGACCCCAATCTGAACCTCTTTGA 1316
Db 721 TAGTAAGGAAACTGCATCAGATGAAAGAGGAAACAGACCCCAATCTGAACCTCTTTGA 780
QY 1317 GTTTACTCATCTGTACAGAGCTGCAGGGAGTGCACAGATGCCAGAGAACTTAGCA 1376
Db 781 GTTTACTCATCTGTACAGAGCTGCAGGGAGTGCACAGATGCCAGAGAACTTAGCA 840
QY 1377 GGGTCTCCCCGGAG-GAGAGGTTTGGGAAGCTCCACGGAGAGGAAAGCTCTCTCTCTCCA 1435
Db 841 GGGTCTCCCCGGAGAGAGAGTTTGGGAGAGTCCACGAGAGGAAAGCTCTCTCTCTCCA 900
QY 1436 GCCTCTTTTCAATGCGGTACAGATGACAGACCTCCAGCATCCAGCATCTCTTGGTCCCA 1495
Db 901 GCCTCTTTTCAATGCGGTACAGATGACAGACCTCCAGCATCCAGCATCAAGGCATCTCTTGGTCCCA 960

```

Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

QY 1496 ATAACTGCTCTAGATACATAGCCATACCTAGTCTAGTAAACCCAGTGTCCCTCAGACTT-GG 1554
 Db 961 ATAACTGCTCTCTA-AAACATAAGCCATCTGCTAGNTAAACCCAGGTCCTCACAACCTGGG 1019
 QY 1555 ATGAGATTCTT-GGAGGGTACACCCAAATGATCAGATACCTTGTATACCTTTTGGAGCCC 1612
 Db 1020 ATGAGATTCTTGGGAGGGAACCCAAATGATGCGCATCTTGTATTAATTTGGAGCCC 1078

RESULT 5
 BM463192 1094 bp mRNA linear EST 05-FEB-2002
 LOCUS AGENCOURT 6447506 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5537917
 DEFINITION 5', mRNA sequence.
 ACCESSION BM463192
 VERSION BM463192.1 GI:18512234
 KEYWORDS EST.
 SOURCE
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1094)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC/DCTD/DTP
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 plate: LLAM12229 row: p column: 14
 High quality sequence stop: 675.
 FEATURES
 source
 1..1094
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5537917"
 /tissue_type="melanotic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_72"
 /note="Organ: skin; Vector: pCMV-Sport6; Site: 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 2 kb. Library constructed by Life
 Technologies."

ORIGIN
 Query Match 26.8%; Score 988.6; DB 12; Length 1094;
 Best Local Similarity 96.0%; Pred. No. 2.9e-189;
 Matches 1031; Conservative 0; Mismatches 41; Indels 2; Gaps 2;

QY 2611 TCGACTTCTGTGATTAATTAACCCCACTCAGTACCTGTTTCAGATGCAATGGATACCA 2670
 Db 1 TCGACTTCTGTGATTAATTAACCCCACTCAGTACCTGTTTCAGATGCAATGGATACCA 60
 QY 2671 AAGATTAATCTTGACATAGATCTCATTTGCGAAGAGCAGATTAAGACCATCAGAAG 2730
 Db 61 AAGATTAAGCCCTTGACATAGATCTCATTTGCGAAGAGCAGATTAAGACCATCAGAAG 120
 QY 2731 GAAATTAATTTAGTGTGTAATGACAGGCACTGTCAGAAACTGTGTGCCAAAATAGAA 2790
 Db 121 GAAATTAATTTAGTGTGTAATGACAGGCACTGTCAGAAACTGTGTGCCAAAATAGAA 180
 QY 2791 TTCCTCTAGTTTTCTTGTCTCTATTGAAGAGAGAAAATTCACCTTTGTAGCATTT 2850
 Db 181 TTCCTCTAGTTTTCTTGTCTCTATTGAAGAGAGAAAATTCACCTTTGTAGCATTT 240
 QY 2851 CAAGCTTTTATGTATCCATCCCATCTTAAAACTCTTCAAACTCCCACTGTTTCAGTCTGAA 2910

Db 241 CAGGCTTTATGTATCCATCCATCTAAATAACTCTTCAAACTCCACTTGTTCACTGAA 300
 QY 2911 ATGCACTCCCTGTCCTCAAGTGCCTTTGGAGAACTCAGAGCAGCGCTTAATCAAAGTT 2970
 Db 301 ATGCACTCCCTGTCCTCAAGTGCCTTTGGAGAACTCAGAGCAGCGCTTAATCAAAGTT 360
 QY 2971 TTACAGAGCCCTTGGACACTATGGGAGGAGGCAAGAGTACACCAATTTGTTAAAGCAAG 3030
 Db 361 TTACAGAGCCCTTGGACACTATGGGAGGAGGCAAGAGTACACCAATTTGTTAAAGCAAG 420
 QY 3031 AAACACAGTGTCTCTTCACTAGTCAATTTAGAAATGTTATCATTCGAAGACTACTCTAC 3090
 Db 421 AAACACAGTGTCTCTTCACTAGTCAATTTAGAAATGTTATCATTCGAAGACTACTCTAC 480
 QY 3091 COTGCAACATTTGAATCCCAAGAGCAATCCACATTTCTCTTGAGTTCTGAGCTTCTGT 3150
 Db 481 COTGCAACATTTGAATCCCAAGAGCAATCCACATTTCTCTTGAGTTCTGAGCTTCTGT 540
 QY 3151 GTAATAGGGCAGTGTCTCTATGCCGTAGAAATCACAATCTGAGGACCATTCATGGA 3210
 Db 541 GTAATAGGGCAGTGTCTCTATGCCGTAGAAATCACAATCTGAGGACCATTCATGGA 600
 QY 3211 AGCTGCTAAATAGCTCTCTGGGAGGCTTCCATAAAGTTTTCATGAGGAGCAACAAC 3270
 Db 601 AGCTGCTAAATAGCTCTCTGGGAGGCTTCCATAAAGTTTTCATGAGGAGCAACAAC 660
 QY 3271 AGGATTAACCTAGTGTCTCTTTCAGCCCTCTAAAGAGCATAGGGCTTAGCTGCAGGC 3330
 Db 661 AGGATTAACCTAGTGTCTCTTTCAGCCCTCTAAAGAGCATAGGGCTTAGCTGCAGGC 720
 QY 3331 TTCCCTGGCTTCTCTCTGCTGTGTAGTTTGTAAACACTATAGCATCTGTTAAGATCCA 3390
 Db 721 TTCCCTGGCTTCTCTCTGCTGTGTAGTTTGTAAACACTATAGCATCTGTTAAGATCCA 780
 QY 3391 GTGTCCATGGAACCTTCCACATGCCGTGACTCTGGAATATACAGTTTGTGGAAGCA 3450
 Db 781 GTGTCCATGGAACCTTCCACATGCCGTGACTCTGGAATATACAGTTTGTGGAAGCA 840
 QY 3451 GGGTTCCTCTGCTCTAACAGCCCGTGGACAGCTGAGTCTGATCTCTTCTTTACACC 3510
 Db 841 GGGTTCCTCTGCTCTAACAGCCCGTGGACAGCTGAGTCTGATCTCTTCTTTACACC 900
 QY 3511 TATGTTTTTAACTAGTCAAACTTCAAGAAACAATCTAAACAGTTTCTGTTGCATATGTG 3570
 Db 901 TATGTTTTTAACTAGTCAAACTTCTCAG-AACCATCTAAACAGTTTCTGTTGCATATGTG 959
 QY 3571 TTGTGAACCTGATTTGATTTAGTAGCTCTCTATATGCA-TTAACTGTTTGTGTA 3629
 Db 960 TTGTGAACCTGATTTGATTTAGTAGCTCTCTATATGCA-TTAACTGTTTGTGTA 1019
 QY 3630 ACTCCTGATTTCTCTTTTCGGATACTATTGATGAATAAGAAATTAAGTGAA 3683
 Db 1020 ACTCCTGATTTCTCTTTTCGGATACTATTGATGAATAAGAAATTAAGTGAA 1073

RESULT 6
 EC006185 1677 bp mRNA linear HTC 17-DEC-2003
 LOCUS Mus musculus follistatin-like, mRNA (cdna clone IMAGE:3498052).
 DEFINITION
 ACCESSION EC006185
 VERSION EC006185.1 GI:13544115
 KEYWORDS HTC.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 1677)
 AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L.,

JOURNAL	12477932
PUBLISHED	2 (pages 1 to 1677)
REFERENCE	Straussburg, R.
AUTHORS	Direct Submission
TITLE	Submitted (02-APR-2001) National Institutes of Health, Mammalian
JOURNAL	

REMARK	NIH-MGC Project URL: http://mgc.ncl.nih.gov
COMMENT	Contact: MGC help desk

Tissue Procurement: Jeffrey Green M.D.,
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amg@bcm.tmc.edu
 Gamarate, F.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louisseged, H.,
 Komar, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nannavati,
 A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.lnl.gov>
 Series: IRAK Plate: 9 Row: k Column: 3
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6679870
 This clone has the following problem: no poly-a-tail.

FEATURES	Location/Qualifiers
source	1. .1677

```
organism="Mus musculus"  
/mgi_type="mRNA"  
/strain="FVB/N"  
/db_xref="taxon:10090"  
/clone_image="taxon:348052"  
/clone_image="Mammary tumor; C3(1) Tag model. Infiltrating  
ductal carcinoma. 5 month old virgin mouse."  
clone_id="NCI_CGAP_Mam6"  
lab_host="DH10B"  
note="Vector: pCMV-Sport6"
```

ORIGIN

Query Match	26.0%;	Score 960.4;	DB 11;	Length 1677;
Best Local Similarity	79.5%;	Pred. No. 1.2e-182;		
Matches 1296;	Conservative	0;	Mismatches 296;	Indels 38; Gaps 12;

[illegible]

Qy	199	GAGCCGGCCGGAAATGTCACATCAACAAGAAAAGGGAAACCACTCTCTCTGCAATTGAGC	251
Db	230	GAGCTGGCAGGGAATGTCGCTCAACAAGAAAGGGAGGCCACGTGCTCTGCAATTGAGC	289
Qy	252	AATGCAAACCTTCAACAAGAGCCTTGTGTGGCACTAAATGGCAAGACTTACCTCAACAT	311
Db	290	AATGCAAACCTTCAACAAGAGCCTTGTGTGGCACTAAATGGCAAGACTTACCTCAACAT	349
Qy	312	GTGAATCTGCATCGAATGCTCCCTCACTGGATCCAAATCCAGTGTGATTACGATGAC	371
Db	350	GTGAATCTGCATCGAATGCTCCCTCACTGGATCCAAATCCAGTGTGATTATGATGGGC	409
Qy	372	ACTGCAGAAAGAAAGAAATCCGTAAGTCCATCTGCAAGCCAGTTGTTTGTATCAATCCA	431
Db	410	ACTGCAGAAAGAAAGAAATCCGTAAGTCCATCTGCAAGCCAGTTGTTCTGTACAGCTA	469
Qy	432	ACCGTATAGATCCGACGTGCACTCAATCCAGTGGCTGGAAGCTGATCATCTCAGATG	491
Db	470	ACCGGATAGATCCGACGGGCGCTCATCCAGTGGCTGGAAGCTGATCATCTCAGATG	529
Qy	492	GCTGTGTTCTTAAAGGCGCACTACAGTGAATCTTACACAGTATTTTAAAGACTTGG	551
Db	530	GCTGTGTTCTTAAAGGCGTACACAGTGAATCTTACACAGTATTTTAAAGACTTGG	589
Qy	552	ATTATGATGATTCTGGCCTGGACTCCAGTGAATTCCTGAAGTTTGTGGAAACGATGA	611
Db	590	ATTATGATGATCTCACTCGATCCAGTGAATTCCTGAAGTTTGTGGAAACGATGA	649
Qy	612	CTGCCATGAATTTTCAACGATCCACACGAGAAACAAACAGTTGCTTATGGGACTCT	671
Db	650	CAGCCATCAACATCCACTTATCAATCAGAGAAACAAACAGTTGCTTATGGGACTCT	709
Qy	672	GTGTGATGCTCTCATTTGAATCTGTGATGAAATGCTGATTGGAACCTCAGCTTCAAG	731
Db	710	GTGTGATGCTCTCATTTGAATCTGTGATGAAATGCTGATTGGAACCTCAGCTTCAAG	769
Qy	732	AGTTTCTCAAGTCTCAACCCATCTTTCAACCTCTGTGAAGAAAGTGTGCTTGGAG	791
Db	770	AGTTTCTCAAGTCTCAACCCATCTTTCAACCTCTGTGAAGAAAGTGTGCTTGGAG	829
Qy	792	ATGAAACGATGACAGATGGAAGCTGAGACGAGGTGACCTGTAACCGCTGTGTGTGCT	851
Db	830	ATGAAACGATGACAGATGGAAGCTGAGACGAGGTGACCTGTAACCGCTGTGTGTGCT	889
Qy	852	GTGGAATTTGGTCTGTACAGCCATGACCTGTGACGGAAGAAATCGAAGGGGGCCACA	911
Db	890	GTGGAATTTGGTCTGTACAGCCATGACCTGTGATGGAAGAAATCGAAGGGGGCTCA	949
Qy	912	CCGAGCAGAGAGGAGATGACCAAGATGTCCAGAGCTCCAAAGACATCAGAAACAG	971
Db	950	CCGAGCAGAGAGGAGATGACCAAGATGTCCAGAGCTCCAAAGACATCAGAAACAG	1009
Qy	972	CTGAAABACCAAGAGGTGAGCAACAAAGATCTTAATGAG-GAGGCAACAACAGTGT	1033
Db	1010	CTGAAABACCAAGAGGTGAGCAACAAAGATCTTAATGAGAGGCAACAAGACCGTGT	1066
Qy	1031	CTGGAATCCAGATCTTCTCACTTCAAGCGCTAGATTCAATATACAAGATGCTGTAC	1090
Db	1070	CTGGAATCCAGATCTTCTCACTTCAAGCGCTAGATTCAATATACAAGATGCTGTAC	1122
Qy	1091	AATGGCGAAATCCAGATTTGCTTAATATAGCAAGATTTATTTGTTATTTGTTT	1155
Db	1126	AATTCACCAATCATAGTAATTTGTTGATGAGCAGCAATCTTAATTTTG---TTTGTT	1187
Qy	1151	TGCATTAAGATATGAAAGTGGCTGCTAGAAAGGAAGGACACAGCTTCATTTCTA	1210
Db	1182	TGCATTAAGAAATGAGGTGGCCACACTAGCAGAGGAAGGCAACACTTCACTGTA	1244
Qy	1211	GGAGAGCTTTAAGAAACGTGAATGATGCTCTGAGGCGCTGAGAGCTTATTAAGAAACGT	1270
Db	1242	GGATGCTTTAAGAAAC--TAAAGACACTTGGAGCAGAGGCAACTTAAGAAACAG	1295

2y	1271	CATCAGATTGAAGAGGAAACAGACCCAAAT----	CTGAACCTCTCTTTGAGTTTACTGCA	1326
Db	1300	CATC-GGGTTGGCAGAGGACAGAGGCAGTTTGAATGAAGCCCTTTCTGGGTCACAGCA	1358	
2y	1327	TCTGT CAGCAGGCTGCAGGAGTGCACACATGCCAGAGAGAACTTAGCAGGTCCTCCC	1386	
Db	1359	GCTGCGAGGAGAAATACAGGAAAAG-----	CATAGAGAAACATTGAACCTAGCCCTG	1408
2y	1387	GGAGGAGAGGTTTGGGAAGCTCCACGGAGGAAACGCTCTCTGCTTCCAGCCTCTTTCCA	1446	
Db	1409	CTGGAGGAAGTGGGGAGCTTTGTAGGAGGAAACCTGCTGCTTT--GACCTTTGCA	1465	
2y	1447	TTGCGGT CAGCATGACAGACCTCAGCATCCACGATCTCTTGCTCCCAATAAC-TGCCT	1505	
Db	1466	CCACTGT CAGCATGACAGAGCTGCAGCAAGTCTGCTCTCTTTGGTCCCAACAATCAC	1525	
2y	1506	CTAGATACATAGCCATCTGCTAGTTAAACCGTGCCTCCAGACTTGGATGAGATTCT	1565	
Db	1526	CTGAACACAGCGCGCCCAACTAGTTA--CCTGTGTCCTCAGCCTTGCATGAGATTCC	1583	
2y	1566	GGGAGGGTACACCCAAATGATGCAGATACTGTATCTTTGAGCCCTTAGCAGCCTAAC	1625	
Db	1584	TGGAGGAGGTGTTAAATGATGACAGACATTATGACTTCAAGCGCCATGGAGACCTAAC	1643	
2y	1626	CAAAATTTAA	1635	
Db	1644	CAAAATTTTA	1653	

RESULT 7	AL570818/c	AL570818	999 bp	linear	EST 31-MAY-2003
LOCUS	AL570818	Homc sapiens	PLACENTA	COT 25-NORMALIZED	Homc sapiens
DEFINITION	clone CSDDI022YC11	3-PRIME	mrna	sequence.	CDNA

511	DB	GATGCTGGAAATTGATGAAGCTCAAGTTGAACAAGATGCTCAGCCTTTATTTTCAGGCTTGA	570
829	QY	GAAGGAGAAAAGGAGACAGACGTAAACGTGTTCAATACCATCCTTACCACCAAGACCTAT	888
571	DB	GAACTTAAATGGGGAGCAGATGAAGAAAGTTTATCACCATCTTTGGAAACACGAAGTGTG	630
889	QY	CCACAACCTTCGAGAGTGTTCCTCAGAAATACACCAAGTACAGTAAGCATGACATGAACAA	948
631	DB	TCTCATTTGAGAAAAGGTGTTTCACAAGTACATGACTATATCAGGATTTCAAATTGAGGAA	690
949	QY	GTTCTGGACCTGGAGTTGAAAGTGACATTGAGAAATGCCTCACAGCTATCGTGAAGTGC	1008
691	DB	ACCATTGACCGGAGACTCTCGCAATTTAGAGCAACTACTCCTTGCTGTGTGAATCT	750
1009	QY	GCCACAAGCAACACCGCTTCTCTTCGAGAGAGCTTCATCAAGCCATGAAAGGTGTGGA	1068
751	DB	ATTGGAAGTATACCTGCCCTACCTTCGAGAGACCTCTATTATGCTATGAAGGAGCTGGG	810
1069	QY	ACTCGCCATAAGGCATTGATCAGGATTTATGGTTTCCCGTTCTCGAAATTGACATGAATGAT	1128
811	DB	ACAGATGATCATACCTCATCAGATCATGGTTCCAGAGTGAGATTCATCTGTTTAACT	870
1129	QY	ATCAAAGCAATTATCAGAAGATGTATGATATCTCCCTTTGCCAAGCCATCTCGGATGAA	1188
871	DB	ATCAGGAAGGAGTTTAGGAAGAAATTTGGCCACTCTCTTTATTCATGATTAAGGAGAT	930
1189	QY	ACCAAGGAGAGTATGAGAAATCTCGTGGCTCTTTGTGGAGGAAA	1235
931	DB	ACATCTGGGGAATATAAGAAAGCTCTTCTGCTGCTCTCCGAGAGA	977

```

RESULT 14
US-09-324-096A-5
; Sequence 5, Application US/09324096A
; Patent No. 6323313
; GENERAL INFORMATION:
; APPLICANT: Tait, Jonathan
; APPLICANT: Brown, David
; TITLE OF INVENTION: ANNEXIN DERIVATIVE WITH ENDOGENOUS CHELATION SITES
; FILE REFERENCE: UOPW-1-13841
; CURRENT APPLICATION NUMBER: US/09/324,096A
; CURRENT FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 981
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(981)
US-09-324-096A-5

Query Match      15.6%; Score 236.6; DB 4; Length 981;
Best Local Similarity 53.1%; Pred. No. 2.5e-62;
Matches 503; Conservative 0; Mismatches 444; Indels 0; Gaps 0;

QY      289  GGTCCGGATCAGCGGTGAGCCCCCTATCTCTACCTTCATCCATCTCTCGATGTCGCTGCC 348
Db      31   GTTCTCAGAGGCACCTGTGACTGACTTCCCTGGATTGTGATGAGCGGGCTGTATGCAGAACT 90

QY      349  TTGCATAAGCCCATATGTTTAAAGTGTGGATGAAGCAACCATCATTCGACATTCCTAACT 408
Db      91   CTTCCGAAGCCTATGAAGGCTTGGGCCACAGATGAGGAGAGCATCTCTGACTCTGTTGACA 150

QY      409  AAGCGAAACAATGSCACAGCGTCAACAGATCAAAGCAGCATATCTCCAGGAAACAGGAAAG 468
Db      151  TCCCGAAGTAATGCTCAGCGCCAGGAAATCTCTCGAGCTTTTAAAGACTCTGTTTGCAGG 210

QY      469  CCCCTGGATGAACACTGAAGAAAGCCCTTACAGGTCACTTTGAGGAGGTTGTTTTAGCT 528
Db      211  GATCTTCTCGATGACCTGAAATCAGAACTAACTGGGAAATTTGAAAAATTTAATTTGTGGCT 270

```

529	QY	CTCCTAAAACCTCCAGCGCAATTTGATGCTGATGAACCTTCGTGCTGCCATGAAGGGCCTT	588
271	Db	CTGATGAAACCCCTCTCGGCTTTATGATGCTTATGAACTGAAACATGCTTTGAAGGGAGCT	330
589	QY	GGAACTGATGAAGATACTCTAAATTTGAGATTTTGGCATCAAGAACTAAACAAGAAATCAGA	648
331	Db	GGAAACAATGAAAAGTAGTCTGACAGAAATTTATTTGCTTCAAGGACACCTGAGAACTGAGA	390
649	QY	GACATTAAACAGGCTCTACAGAGAGGAACCTGAAGAGAGATCTGGCCAAAGACATAAACCCTCA	708
391	Db	GCCATCAAAACAAGTTTATGAAGAAGAATATGGCTCTAGCCCTGGAAGATGACGTGTGGGG	450
709	QY	GACACATCTGGAGATTTTCGGAACGCTTTGCTTCTCTTCTTAAGGGTGACCCGATCTGAG	768
451	Db	GACACITTCAGGTACTACACGCGATGTTGGTGGTCTCTCTTCAGGCTAACAGAGCCCT	510
769	QY	GACITTTGGTGTGAATGAAGACTTTGGCTGATTCAGATGCCAGGCGCTTGTATGAAGCAGGA	828
511	Db	GATGCTGGAATTTGATGAAGCTCAAGTTTGAACAAGATGCTCAGGCTTTATTTTCAGGCTGGA	570
829	QY	GAAGAGGAAAGGGGACAGAGCTTAAACGTCTTCAATACCATCTCTTACCACCAAGAGCTAT	888
571	Db	GAACCTTAAATGGGGGACAGATGAAGAAAGTTTATCACCATCTTTTGGAAACAGAGTGTC	630
889	QY	CCACAACTTCGCGAGAGTGTTTTCAGAAATACACCAAGTACAGTAAGCATGACATGAAACAAA	948
631	Db	TCTCATTTGAGAAGGTGTTTGTGCAAGTACATGACTATATCAGGATTTCAAATTGAGGAA	690
949	QY	GTTCCTGCACTGAGTTGAAAGGTGACATTTGAGAAATGCCCTCACAGCTATCGTGAAGTGC	1008
691	Db	ACCATTCACGCGAGACTTCTGGCAATTTAGAGCAACTACTCCTTGTGTTGTGAAATCT	750
1009	QY	GCCACAAGCAAAACCAAGCTTTCTTTTCAGAGAAAGCTTCATCAAGCCATCAAGGTGTGGA	1068
751	Db	ATTCGAAGTATACCTGCTACTTTGCAGAGACCTCTATTATGCTATGAAGGAGCTGGG	810
1069	QY	ACTGCGCATTAAGGCATTGATCAGGATTTATGGTTTCCGTTCTCGAAATTCATGAATGAT	1128
811	Db	ACAGATGATCATACCCCTCATCAGAGTCAATGGTTTCCAGGAGTGAGATTGATCTCTTTTAA	870
1129	QY	ATCAAGCAATTCATCAGAGAGATGATGGTATCTCCCTTTTGCCAAAGCCATCTCGATGAA	1188
871	Db	ATCAGGAAGAGTTTAGAAGAAATTTTGGCCACCTCTCTTTATTCCTATGATTAAGGGAGAT	930
1189	QY	ACCAAAGGAGTAGTGAAGAAATCCTGGTGGCTCTTTGTGGAGAA	1235
931	Db	ACATCTGGGACTATAGAAGAGCTTCTGCTGCTCTCCGAGAGA	977

RESULT 15
US-09-976-594-348
Sequence 348, Application US/09976594
Patent No. 6673549
GENERAL INFORMATION:
APPLICANT: Furness, Michael
APPLICANT: Buchbinder, Jenny
TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
FILE REFERENCE: PA-0041 US
CURRENT APPLICATION NUMBER: US/09/976,594
CURRENT FILING DATE: 2001-10-12
PRIOR APPLICATION NUMBER: 60/240,409
PRIOR FILING DATE: 2000-10-12
NUMBER OF SEQ ID NOS: 1143
SOFTWARE: PERL Program
SEQ ID NO 348
LENGTH: 2137
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. 6673549 043646.7
US-09-976-594-348

	Query Match	15.4%;	Score 232.8;	DB 4;	Length 2137;
	Best Local Similarity	52.8%;	Pred. No. 6.1e-61;		
	Matches 501;	Conservative	0;	Mismatches 447;	Indels 0; Gaps 0;
QY	308	GOCCCTATCCTACCTTCAATCCATCCTCGGATGTGCTGCTGCCCTTGCATATAGCCGCAATATGG	367		
DB	545	GACCAGCTGCCAACTTCGATGCTATAAGAGATGCGAAATTTCTTCGTAAGGCCAATGAAGG	604		
QY	368	TTAAAGGTGTGGATGAAGCAACCATCATTTGACATTTCTAACTAAGACCGAAACAATGSCACAGC	427		
DB	605	GTTTTGGGACAGATGAGCAGGCAATTTGTGGATGTGTGGCGCAACCGTTCCATGATCAGA	664		
QY	428	GTCAACAGATCAAAAGCAGCATATCTCCAGAAACAGAAAGCCCTGGATGAAACACTGA	487		
DB	665	GGCAAAAAATTAAAGCAGCATTTAAGACCTCCTATGGCAAGGATTTAAATCAAAAGATCTCA	724		
QY	488	AGAAAGCCCTTACAGCTACCTTGAGAGGTTGTTTTAGCTCTGCTATAAAACTCCAGCGC	547		
DB	725	AATCAGAGTTAAGTGAATATATGGAAGAACTGATCTCTGGCCCTCTTCATGCCCTCTACGT	784		
QY	548	AAATTGATGCTGATGAACCTTCTGCTGCCATGAAGGCCCTTGGAACTGATGAAGATACTC	607		
DB	785	ATTAACGATCCCTGGAGCTTAACGAAAGCAATGCAGGAGCAGGAACTCAGAAACGTTAT	844		
QY	608	TAATTGAGATTTTGGCATCMAGAACTTAAACAGAAATCAGAGACATTAACAGGGTCTACA	667		
DB	845	TGATTGAGATTTTGTGCCAAGAAACAAATCAGAAATCCGAGAAATTTGTGAGATGTTATC	904		
QY	668	GAGAGGAACCTGAAGACAGATCTCGCCAAAGACATAACCTCAGACACATCTCGAGATTTTC	727		
DB	905	AGTCAGAAATTGGACGAGACCTTTGAAAGGACATTAGGTCAGATACATCAGGACATTTTG	964		
QY	728	GGAAACGTTTGTCTTCTCTTAAGGGTGACCGATCTGAGGACTTTGGTGTGAATGAAG	787		
DB	965	AACGTTTACTGTGTCCATGTGCCAGGGAATCGTGATGAGAACCGAGATATAACCCACC	1024		
QY	788	ACTTGGCTGATCAGATGCCAGGCGCTTGTATGAAGCAGGAGAAAGAGAGAAAGGGGACAG	847		
DB	1025	AAATGGCTCAGGAAGATGCTCAGCGTCTCTATCAAGCTGTGTAGGGGAGACTAGGGACCG	1084		
QY	848	ACGTAAACGTGTTCAATACCATCTTACACAGAGACTATCCACAACCTTCGCAGAGTGT	907		
DB	1085	ATGAATCTTGCTTTAATCATGATCCTTGCCCAAGAAGCTTTCTCTACGTGAGAGCTACCA	1144		
QY	908	TTCAGAAATACACCAAGTACAGTAAGCATGACATGAACAAAGTTCTGGACCTGGAGTTGA	967		
DB	1145	TGAGGCTTATTTCTAGGATGGCTAATCGAGCTTGTTAAGCAGTGTGAGCGCTGAGTTTT	1204		
QY	968	AAGTGACATGAGAAATGCGCTCACAGCTATCGTGAAGTGGCCCAAGACMAACAGCTT	1027		
DB	1205	CCGGATATGTAGAAAGTGGTTTGAAGACCACTCTTGCASTGTGCCCTGAACCGCGCTGCCT	1264		
QY	1028	TCCTTTCAGAGAGCTTCATCAAGCCATGAAAGGTGTGGAACCTCGGCCATAAAGCAATGA	1087		
DB	1265	TCCTTTCGAGAGGCTCTACTATGCTATGAAAGGTGTGSCACAGATGACTCCACCCTGG	1324		
QY	1088	TCAGGATTAATGGTTTCCGTTCTGAAATTTGACATGAATGATATCAAAGCATTTCTATCAGA	1147		
DB	1325	TCCGATTTGTGGTCACTCGAAGTGAGATTGACCTTGTACAAATAAAAACAGATGTTCCGCTC	1384		
QY	1148	AGATGATGATCTCTCCCTTTTGGCAAGCCCATCTGGATGAAACGCGCCATAAAGCAATGA	1207		
DB	1385	AGATGATCAGAGACTCTGGGCACATGATGTCAGGTGACACGAGTGGAGATTACCGAA	1444		
QY	1208	AAATCCCTGGTGGCTTTTGTGGAGGAAACTAAACATTCCTCCCTTGTAGTGT	1255		
DB	1445	GACTTCTCTCGCTATTGTGGCGCAGTAGAGGGGATTTTTTTTTTTTGT	1492		

Search completed: March 11, 2004, 14:41:25
Job time : 72.2294 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

2M nucleic - nucleic search, using sw model

run on: March 11, 2004, 03:37:21 ; Search time 353.257 Seconds
(without alignments)
15799.545 Million cell updates/sec

Title: US-10-084-817-8

Perfect score: 1516

Sequence: 1 cttgtttttggacatagct.....gctgaataataatgncgtrc 1516

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2432557 seqs, 1840798884 residues

Total number of hits satisfying chosen parameters: 4865114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgm2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgm2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgm2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgm2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgm2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgm2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgm2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgm2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgm2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgm2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgm2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgm2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgm2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgm2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgm2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgm2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 17: /cgm2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 18: /cgm2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1515	99.9	1516	9	US-09-919-172-38
2	1515	99.9	1516	9	US-09-974-288-92
3	1515	99.9	1516	14	US-10-084-817-8
4	1389.4	91.6	1399	9	US-09-864-864-321
5	1387.8	91.5	1399	9	US-09-954-456-2119
6	1387.8	91.5	1399	14	US-10-240-965-2
7	1387.8	91.5	1399	14	US-10-269-909-2
8	1387.8	91.5	1399	15	US-10-133-937-2
9	1387.8	91.5	1399	15	US-10-159-563-2
10	1379.4	91.0	1431	9	US-09-925-301-46
11	1367.4	90.2	1377	14	US-10-210-120-18
12	886.8	58.5	1402	9	US-09-917-800A-1585
13	886.8	58.5	1402	14	US-10-316-253-299
14	832	54.2	2500	10	US-09-814-353-20447
15	597.4	39.4	600	12	US-10-085-783A-51438

16	597.4	39.4	600	15	US-10-242-535A-51438	Sequence 51438, A
17	589.4	38.9	614	9	US-09-864-864-321	Sequence 241, App
18	583.4	38.5	618	12	US-10-085-783A-13898	Sequence 13898, A
19	583.4	38.5	618	15	US-10-242-535A-13898	Sequence 13898, A
20	529	34.9	592	12	US-10-085-783A-48352	Sequence 48352, A
21	529	34.9	592	15	US-10-242-535A-48352	Sequence 48352, A
c	519.4	34.3	537	9	US-09-777-584-134	Sequence 134, App
c	519.4	34.3	537	14	US-10-015-219-134	Sequence 134, App
22	505	33.3	557	12	US-10-085-783A-48313	Sequence 48313, A
23	505	33.3	557	15	US-10-242-535A-48313	Sequence 48313, A
24	500	33.0	530	12	US-10-085-783A-54108	Sequence 54108, A
25	500	33.0	530	15	US-10-242-535A-54108	Sequence 54108, A
26	500	33.0	530	12	US-10-085-783A-48227	Sequence 48227, A
27	496	32.7	590	12	US-10-085-783A-48227	Sequence 48227, A
28	496	32.7	590	15	US-10-242-535A-48227	Sequence 48227, A
29	491.4	32.4	500	12	US-10-085-783A-23999	Sequence 23999, A
30	491.4	32.4	500	15	US-10-242-535A-23999	Sequence 23999, A
31	482	31.8	495	12	US-10-085-783A-38267	Sequence 38267, A
32	482	31.8	495	15	US-10-242-535A-38267	Sequence 38267, A
33	474	31.3	499	12	US-10-085-783A-26128	Sequence 26128, A
34	474	31.3	499	15	US-10-242-535A-26128	Sequence 26128, A
35	468.2	30.9	473	12	US-10-085-783A-24370	Sequence 24370, A
36	468.2	30.9	473	15	US-10-242-535A-24370	Sequence 24370, A
37	460.2	30.4	470	12	US-10-085-783A-42942	Sequence 42942, A
38	460.2	30.4	470	15	US-10-242-535A-42942	Sequence 42942, A
39	459.4	30.3	461	9	US-09-998-598-607	Sequence 607, App
40	458	30.2	473	12	US-10-085-783A-31104	Sequence 31104, A
41	458	30.2	473	15	US-10-242-535A-31104	Sequence 31104, A
42	457.8	30.2	461	9	US-09-998-598-1129	Sequence 1129, App
43	455.2	30.0	472	12	US-10-085-783A-55749	Sequence 55749, A
44	455.2	30.0	472	15	US-10-242-535A-55749	Sequence 55749, A
45	455.2	30.0	472	15	US-10-242-535A-55749	Sequence 55749, A

ALIGNMENTS

RESULT 1

US-09-919-172-38

; Sequence 38, Application US/09919172

; Patent No. US20020119463A1

; GENERAL INFORMATION:

; APPLICANT: Faris, Mary

; APPLICANT: Turner, Christopher M.

; TITLE OF INVENTION: PROSTATE CANCER MARKERS

; FILE REFERENCE: PA-0036 US

; CURRENT APPLICATION NUMBER: US/09/919,172

; CURRENT FILING DATE: 2001-07-30

; PRIOR APPLICATION NUMBER: 60/222,469

; PRIOR FILING DATE: 2000-07-28

; NUMBER OF SEQ ID NOS: 102

; SOFTWARE: PERL Program

; SEQ ID NO 38

; LENGTH: 1516

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; OTHER INFORMATION: Incyte ID No. US20020119463A1 1303785CB1

; NAME/KEY: unsure

; LOCATION: 1512

; OTHER INFORMATION: a, t, c, g, or other

US-09-919-172-38

Query Match 99.9%; Score 1515; DB 9; Length 1516;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTGTGTTTGGACATAGCTGAGCCATGCTACTTCAACAGAGGCGCAATTAACCT 60

Db 1 CTTGTGTTTGGACATAGCTGAGCCATGCTACTTCAACAGAGGCGCAATTAACCT 60

QY 61 TCTGGTGTCTAGCTGGCTTCTTTTAAATCTTATAAATCAGAGCCCACTCTCCAC 120

Db 61 TCTGGTGTCTAGCTGGCTTCTTTTAAATCTTATAAATCAGAGCCCACTCTCCAC 120

241 GAAATGAAGAGCAGGAATATGTTCAAACTGTAAGTCAATCCAAAGGTGGTCCCGATCA 300
301 GCGGTGAGCCCTTATCTCTACCTTCAATCCATCTCGATGTCGCTCTTGCATAGGCC 360
301 GCGGTGAGCCCTTATCTCTACCTTCAATCCATCTCGATGTCGCTCTTGCATAGGCC 360
361 ATAATGGTTAAAGGTGTGGATGAAGCAACCATCATTTGACATTTAACTAAGCGAAACAT 420
361 ATAATGGTTAAAGGTGTGGATGAAGCAACCATCATTTGACATTTAACTAAGCGAAACAT 420
421 GCACAGGCTCAACAGATCAAGCAGATATCTCCAGGAAACAGGAAAGCCCTGGATGAA 480
421 GCACAGGCTCAACAGATCAAGCAGATATCTCCAGGAAACAGGAAAGCCCTGGATGAA 480
481 ACACGAGAAAGCCCTTACAGGTCACTTGGAGAGGTGTTTGTAGCTGTCTGTAACACT 540
481 ACACGAGAAAGCCCTTACAGGTCACTTGGAGAGGTGTTTGTAGCTGTCTGTAACACT 540
541 CCAGCGCAATTTGATGCTGATGAATCGTCTGCTGCCATGAAGGCCCTTGAACATGAA 600
541 CCAGCGCAATTTGATGCTGATGAATCGTCTGCTGCCATGAAGGCCCTTGAACATGAA 600
601 GATACCTTAATTTGAGATTTTGGCATCAAGAACTAAACAAAGAAATCAGACATTAACAG 660
601 GATACCTTAATTTGAGATTTTGGCATCAAGAACTAAACAAAGAAATCAGACATTAACAG 660
661 GTCTACAGAGAGAACTGAAGAGATCTGGCCAAAGACATTAACCTCAGACATCTGGA 720
661 GTCTACAGAGAGAACTGAAGAGATCTGGCCAAAGACATTAACCTCAGACATCTGGA 720
721 GATTTTCGGAACGCTTTGCTTCTTCTGCTAAGGAGTACCGATCTGAGGACTTTGGTGTG 780
721 GATTTTCGGAACGCTTTGCTTCTTCTGCTAAGGAGTACCGATCTGAGGACTTTGGTGTG 780
781 AATGAAGACTTGGTGTGATTCAGATGCGCAGGCGCTTGTATGAAGCAGAGAAAGGAAAG 840
781 AATGAAGACTTGGTGTGATTCAGATGCGCAGGCGCTTGTATGAAGCAGAGAAAGGAAAG 840
841 GGGACAGAGCTGAACGCTTTCAATACCATCTCTACCAAGAGAGCTATCACAACCTTCGC 900
841 GGGACAGAGCTGAACGCTTTCAATACCATCTCTACCAAGAGAGCTATCACAACCTTCGC 900
901 AGAGTGTTCAGAAATACACAAATGACATGATGAAGATGACATGAACAAAGTCTCGACCTG 960
901 AGAGTGTTCAGAAATACACAAATGACATGATGAAGATGACATGAACAAAGTCTCGACCTG 960
961 GAGTTGAAGGTGACATTTGAGAAATGCTTCAAGCTATCGTGAAGTGGCGCAACAGCAA 1020
961 GAGTTGAAGGTGACATTTGAGAAATGCTTCAAGCTATCGTGAAGTGGCGCAACAGCAA 1020
1021 CCAGCTTTCTTTGACAGAGAGCTTCAACAGCCATGAAGGTGTTGAACTCGCCATPAAG 1080
1021 CCAGCTTTCTTTGACAGAGAGCTTCAACAGCCATGAAGGTGTTGAACTCGCCATPAAG 1080
1081 GCATTGATCAGGATTTAGTTTCCGTTCTGAATTCGATGATGATATCAAGACATTC 1140
1081 GCATTGATCAGGATTTAGTTTCCGTTCTGAATTCGATGATGATATCAAGACATTC 1140
1141 TATCAGAAAGATGATGTTGATCTCCCTTTGCCAAGCCATCTCGGATGAAACCAAGGAGAG 1200
1141 TATCAGAAAGATGATGTTGATCTCCCTTTGCCAAGCCATCTCGGATGAAACCAAGGAGAG 1200
1201 TATGAGAAATCTCGTGGTCTTTGAGGAGAACTTAAACATTTCCCTTGTGATGCTCAA 1260
1201 TATGAGAAATCTCGTGGTCTTTGAGGAGAACTTAAACATTTCCCTTGTGATGCTCAA 1260
1261 GCTATGATCAGAAAGCTTTAATATATATTTTATCTTATAGCTTAAATAGGAAAGTTT 1320
1261 GCTATGATCAGAAAGCTTTAATATATATTTTATCTTATAGCTTAAATAGGAAAGTTT 1320
1321 CTTCAACAGGATTAAGGTAGCTACCTATACGCTGGAATAATATAGCCCTTTAAATCATTT 1380
1321 CTTCAACAGGATTAAGGTAGCTACCTATACGCTGGAATAATATAGCCCTTTAAATCATTT 1380

QY 1381 TTATATTATAACTCTGTATATAATAGATAAAGTCAATTTTAAAAATGTTTTCCTCCAAAC 1440
Db 1381 TTATATTATAACTCTGTATATAATAGATAAAGTCAATTTTAAAAATGTTTTCCTCCAAAC 1440
QY 1441 CATAAACCCCTATACAAGTTGTTCTAGTAAACAATACATGAGAAAGATGTTCTATAGTCTG 1500
Db 1441 CATAAACCCCTATACAAGTTGTTCTAGTAAACAATACATGAGAAAGATGTTCTATGCTG 1500
QY 1501 AAAATAAAATGNCGTC 1516
Db 1501 AAAATAAAATGNCGTC 1516

RESULT 3
US-10-084-817-8
; Sequence 8, Application US/10084817
; Publication No. US20030119009A1
; GENERAL INFORMATION:
; APPLICANT: Susan Stuart
; APPLICANT: Jed G. Nuchtern
; APPLICANT: Sharon E. Plon
; APPLICANT: Jason M. Shohet
; TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION
; FILE REFERENCE: PA-0046 US
; CURRENT APPLICATION NUMBER: US/10/084,817
; PRIOR FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 60/270,784
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 365
; SOFTWARE: PERL Program
; SEQ ID NO 8
; LENGTH: 1516
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030119009A1 1303785CB1
; NAME/KEY: unsure
; LOCATION: 1512
; OTHER INFORMATION: a, t, c, g, or other
US-10-084-817-8

Query Match 99.9%; Score 1515; DB 14; Length 1516;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTTTGTGTTTGGACATAGCTGAGCCATGTACTTCAACAGAGGCGAGCCAAATTACTAACT 60
Db 1 CTTTGTGTTTGGACATAGCTGAGCCATGTACTTCAACAGAGGCGAGCCAAATTACTAACT 60
QY 61 TGTGTTGTAGTGTGGCTTCTTAAATCCTATAAATCAGAGCCAGTCTCCAC 120
Db 61 TGTGTTGTAGTGTGGCTTCTTAAATCCTATAAATCAGAGCCAGTCTCCAC 120
QY 121 TGCCAGTGTGAATCTTCAGAGAAGAAATTTCTTTAGTTCTTTTCAAGAAGGTAGAGAT 180
Db 121 TGCCAGTGTGAATCTTCAGAGAAGAAATTTCTTTAGTTCTTTTCAAGAAGGTAGAGAT 180
QY 181 AAGACACATTTTCAAAAATGCGATGATCAGAAATTCCTCAAGCAGCCCTGTTTAT 240
Db 181 AAGACACATTTTCAAAAATGCGATGATCAGAAATTCCTCAAGCAGCCCTGTTTAT 240
QY 241 GAAATGAAGAGCAGGAATATGTTCAAACTGTGAAGTCAATCCAAAGGTGGTCCCGATCA 300
Db 241 GAAATGAAGAGCAGGAATATGTTCAAACTGTGAAGTCAATCCAAAGGTGGTCCCGATCA 300
QY 301 GCGGTGAGCCCTTATCTCTACCTTCAATTCATCTCGATGTCGCTGCCCTTGCATAGGCC 360
Db 301 GCGGTGAGCCCTTATCTCTACCTTCAATTCATCTCGATGTCGCTGCCCTTGCATAGGCC 360
QY 361 ATAATGTTTAAAGGTGTGATGAAGCAACCATCATTTGACATTTCTAACTAAGCGAAACAT 420

Db 361 ATAAATGGTTAAAGGTGTGGATGAAGCAACCATCATTCATGACATTTCTAAGCGAAACAAT 420
QY 421 GCACAGCGTCAACAGATCAAGACGACATATCTCCAGGAACAGGAAGCCCTGGATGAA 480
Db 421 GCACAGCGTCAACAGATCAAGACGACATATCTCCAGGAACAGGAAGCCCTGGATGAA 480
QY 481 ACATCTGAAGAAGCCCTTACAGGTCACTTGGAGAGTGTGTTTAGCTCTGCTAAAAACT 540
Db 481 ACATCTGAAGAAGCCCTTACAGGTCACTTGGAGAGTGTGTTTAGCTCTGCTAAAAACT 540
QY 541 CAGGCGCAATTGATGCTGATGACACTTGGTCCCATGAGGCGCTTGGAACTGATGAA 600
Db 541 CAGGCGCAATTGATGCTGATGACACTTGGTCCCATGAGGCGCTTGGAACTGATGAA 600
QY 601 GATACCTCTAAATGAGATTTGGCATCAAGAACTAACAAAGAAATCAGAGACATTAACAGG 660
Db 601 GATACCTCTAAATGAGATTTGGCATCAAGAACTAACAAAGAAATCAGAGACATTAACAGG 660
QY 661 GTCTACAGAGAGAACTGAAGAGAGATCTGCCCAAGACATACCTCAGACACATCTGGA 720
Db 661 GTCTACAGAGAGAACTGAAGAGAGATCTGCCCAAGACATACCTCAGACACATCTGGA 720
QY 721 GATTTTCGGAACGCTTGTCTTCTTCTTAAGGTGACCGATCTGAGGACTTTGGTGTG 780
Db 721 GATTTTCGGAACGCTTGTCTTCTTCTTAAGGTGACCGATCTGAGGACTTTGGTGTG 780
QY 781 AATGAAGACTTGGCTGATTCAGATCCGAGGCGCTTGTATGAAGCAGAGAAAGAGAAAG 840
Db 781 AATGAAGACTTGGCTGATTCAGATCCGAGGCGCTTGTATGAAGCAGAGAAAGAGAAAG 840
QY 841 GGGACAGAGCTAAACGCTTCAATACCATCTTACCACCAAGCTATCCACAACTTCGC 900
Db 841 GGGACAGAGCTAAACGCTTCAATACCATCTTACCACCAAGCTATCCACAACTTCGC 900
QY 901 AGAGTGTTCAGAAATACACCAAGTACAGTAAAGTACATGAACAAAGTTCTGGACCTG 960
Db 901 AGAGTGTTCAGAAATACACCAAGTACAGTAAAGTACATGAACAAAGTTCTGGACCTG 960
QY 961 GAGTTGAAAGGTGACATGAGAAATAGCCTCAGCTATCGTGAAGTGGCCACAGCBA 1020
Db 961 GAGTTGAAAGGTGACATGAGAAATAGCCTCAGCTATCGTGAAGTGGCCACAGCBA 1020
QY 1021 CCAGCTTTCTTTGCAGAGAAGCTTCATCAAGCCATGAAAGGTGTTGGAACCTCGCCATAAG 1080
Db 1021 CCAGCTTTCTTTGCAGAGAAGCTTCATCAAGCCATGAAAGGTGTTGGAACCTCGCCATAAG 1080
QY 1081 GCATTGATCAGGATTAATGTTTCCGCTTCTGAAATGACATGAATGATATCAAGCATTC 1140
Db 1081 GCATTGATCAGGATTAATGTTTCCGCTTCTGAAATGACATGAATGATATCAAGCATTC 1140
QY 1141 TATCAGAAGATGATGTTATCTCCCTTTGCGCAAGCCATCTCGATGAAACCAAGAGAG 1200
Db 1141 TATCAGAAGATGATGTTATCTCCCTTTGCGCAAGCCATCTCGATGAAACCAAGAGAG 1200
QY 1201 TATGAGAAATCTGCTGCTCTTTGAGGAAATCAACATCCCTTGATGGTCTCAA 1260
Db 1201 TATGAGAAATCTGCTGCTCTTTGAGGAAATCAACATCCCTTGATGGTCTCAA 1260
QY 1261 GCTATGATCAGAAGACTTTAAATATATATTTTATATTTTATATATTTTAAATAGAAAGTTT 1320
Db 1261 GCTATGATCAGAAGACTTTAAATATATATTTTATATTTTATATATTTTAAATAGAAAGTTT 1320
QY 1321 CTTCAACAGGATTAAGTGTAGTACTACATGCTGAAATAATATAGCCTTTAAATCATTT 1380
Db 1321 CTTCAACAGGATTAAGTGTAGTACTACATGCTGAAATAATATAGCCTTTAAATCATTT 1380
QY 1381 TTATATTATAACTCTGTATAATAGAGATAAGTCCATTTTTTAAATGTTTTTCCCCAAC 1440
Db 1381 TTATATTATAACTCTGTATAATAGAGATAAGTCCATTTTTTAAATGTTTTTCCCCAAC 1440
QY 1441 CATAAAACCCCTATACAGTTGTTCTAGTAAACATATACATGAGAAAGTGTCTATGAGCTG 1500
Db 1441 CATAAAACCCCTATACAGTTGTTCTAGTAAACATATACATGAGAAAGTGTCTATGAGCTG 1500

QY 1501 AAAATAAAATGNCGTC 1516
Db 1501 AAAATAAAATGNCGTC 1516

RESULT 4
US-09-864-864-321
; Sequence 321, Application US/09864864
; Patent No. US20020102679A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Dillon, Devin C.
; APPLICANT: Secrist, Heather J.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Algate, Paul A.
; APPLICANT: Fling, Steve P.
; APPLICANT: Mannion, Jane
; APPLICANT: Benson, Darin R.
; APPLICANT: Carter, Darick
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.523
; CURRENT APPLICATION NUMBER: US/09/864,864
; CURRENT FILING DATE: 2001-05-23
; NUMBER OF SEQ ID NOS: 341
; SOFTWARE: Corixa Invention Disclosure Database
; SEQ ID NO 321
; LENGTH: 1399
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-864-864-321

Query Match 91.6%; Score 1389.4; DB 9; Length 1399;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1390; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 125 AGTGTGAATCTTTCAGAGAAGAAATTTCTTTAGTCTTTTTCAGAGAAGTGTAGATGAAAG 184
Db 1 AGTGTGAATCTTTCAGAGAAGAAATTTCTTTAGTCTTTTTCAGAGAAGTGTAGATGAAAG 60
QY 185 ACATTTTTCAGAAATGCAATGGTATCAGATTCCTCAAGCAGGCTGTTTATTGAAA 244
Db 61 ACATTTTTCAGAAATGCAATGGTATCAGATTCCTCAAGCAGGCTGTTTATTGAAA 120
QY 245 ATGAAGACGAGGAATATGTTTCAAACTGTGAAGTCATCCAAAGGTGGTCCCGATCAGCG 304
Db 121 ATGAAGACGAGGAATATGTTTCAAACTGTGAAGTCATCCAAAGGTGGTCCCGATCAGCG 180
QY 305 TGAGCCCTTATCTTACCTTCATCCATCCGATGCTCGATGCTGCTTGCATTAAGGCCATAA 364
Db 181 TGAGCCCTTATCTTACCTTCATCCATCCGATGCTCGATGCTGCTTGCATTAAGGCCATAA 240
QY 365 TGGTTAAAGGTGGGATGAAGCAACCATCATTCACATTTCTAACTAAGCGAAACAATGCAC 424
Db 241 TGGTTAAAGGTGGGATGAAGCAACCATCATTCACATTTCTAACTAAGCGAAACAATGCAC 300
QY 425 AGCGTCAACAGATCAAGCAGCATATCTCCAGAAACAGGAAGCCCTGGATGAAACAC 484
Db 301 AGCGTCAACAGATCAAGCAGCATATCTCCAGAAACAGGAAGCCCTGGATGAAACAC 360
QY 485 TGAAGAAGCCCTTACAGGTCACTTGGAGAGTGTGTTTTAGTCTCTGCTAAAACTCCAG 544
Db 361 TGAAGAAGCCCTTACAGGTCACTTGGAGAGTGTGTTTTAGTCTCTGCTAAAACTCCAG 420
QY 545 CGCAATTTGATGCTGATGAATCTTCTGCTGCCATGAAGGCCCTTGGAACTGATGAAGATA 604
Db 421 CGCAATTTGATGCTGATGAATCTTCTGCTGCCATGAAGGCCCTTGGAACTGATGAAGATA 480
QY 605 CTCTAAATTTGAGATTTTGGCATCAAGAACTAAAGAAATCAGAGACATTAACAGGGTCT 664

QY 965 TGAAGGTGACATTCAGAAATGCTTCACAGCTATCGTGAAGTGGCCACAAAGCAACCCAG 1024
 DB 841 TGAAGGTGACATTCAGAAATGCTTCACAGCTATCGTGAAGTGGCCACAAAGCAACCCAG 900
 QY 1025 CTTTCTTTGAGAGAGAGCTTCATCAAGCCATGAAGGTGTTGGAACTGCGCATAGGCAT 1084
 DB 901 CTTTCTTTGAGAGAGAGCTTCATCAAGCCATGAAGGTGTTGGAACTGCGCATAGGCAT 960
 QY 1085 TGATCAGGATATGTTTCCCGTTCGAAATGACATGAATGATATCAAAAGCAATTTCTATC 1144
 DB 961 TGATCAGGATATGTTTCCCGTTCGAAATGACATGAATGATATCAAAAGCAATTTCTATC 1020
 QY 1145 AGAAGATGATGATATCTCCCTTTGCGAAGCCATCCTGGATGAACCAAGAGAGATATG 1204
 DB 1021 AGAAGATGATGATATCTCCCTTTGCGAAGCCATCCTGGATGAACCAAGAGAGATATG 1080
 QY 1205 AGAAATCTGCTGCTCTTTGCGAAGAACTAAACATTCCTTGTGTTCTCAAGCTA 1264
 DB 1081 AGAAATCTGCTGCTCTTTGCGAAGAACTAAACATTCCTTGTGTTCTCAAGCTA 1140
 QY 1265 TGATCAGAGACTTTAAATATATATTTTCATCCCTATAGCTTAATAGGAAAGTTCTTC 1324
 DB 1141 TGATCAGAGACTTTAAATATATATTTTCATCCCTATAGCTTAATAGGAAAGTTCTTC 1200
 QY 1325 RACAGGATACAGTGTAGCTACCTACATGCTGAAATATATAGCTTTAAATCAATTTTAT 1384
 DB 1201 RACAGGATACAGTGTAGCTACCTACATGCTGAAATATATAGCTTTAAATCAATTTTAT 1260
 QY 1385 ATTATAACTCTGTATAATAGAGATAAGTCCATTTTAAAAAATGTTTCCCAAAACCAT 1444
 DB 1261 ATTATAACTCTGTATAATAGAGATAAGTCCATTTTAAAAAATGTTTCCCAAAACCAT 1320
 QY 1445 AAACCTATACAGTGTCTAGTAACTATACATGAGAGAGATCTCTATGATGAGTGA 1504
 DB 1321 AAACCTATACAGTGTCTAGTAACTATACATGAGAGAGATCTCTATGATGAGTGA 1380
 QY 1505 TAAATGNCCTC 1516
 DB 1381 TAAATGAGCTC 1392

RESULT 7
 US-10-269-909-2
 ; Sequence 2, Application US/10269909
 ; Publication No. US20030180747A1
 ; GENERAL INFORMATION:
 ; APPLICANT: HRUBAN, RALEH H.
 ; APPLICANT: ARGANI, PEDRAM
 ; APPLICANT: JACOBUIZIO-DONAHUE, CHRISTINE
 ; APPLICANT: MAITRA, ANIRBAN
 ; TITLE OF INVENTION: PANCREATIC CANCER DIAGNOSIS AND THERAPIES
 ; FILE REFERENCE: 58303(71699)
 ; CURRENT APPLICATION NUMBER: US/10/269,909
 ; CURRENT FILING DATE: 2003-10-11
 ; PRIOR APPLICATION NUMBER: 60/328,609
 ; PRIOR FILING DATE: 2001-10-11
 ; PRIOR APPLICATION NUMBER: 60/332,754
 ; PRIOR FILING DATE: 2001-11-19
 ; NUMBER OF SEQ ID NOS: 87
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 1399
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-269-909-2

Query Match 91.5%; Score 1387.8; DB 14; Length 1399;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1389; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 125 AGTGTGAATCTTCAGAGAGAAATTTCTTTTAGTCTTTTGAAGGTAGAGATAAAG 184

DB 1 AGTGTGAATCTTCAGAGAGAAATTTCTTTTAGTCTTTTGAAGGTAGAGATAAAG 60
 QY 185 ACACATTTTCAAAAATGGCAATGGTATCAGAAATCCTCAAGCAGGCCCTGGTTTATGAAA 244
 DB 61 ACACATTTTCAAAAATGGCAATGGTATCAGAAATCCTCAAGCAGGCCCTGGTTTATGAAA 120
 QY 245 ATGAGAGAGAGAAATATGTTCAAACTGTGAAGTCAATCAAAAGGTGGTCCCGATCAGCGG 304
 DB 121 ATGAGAGAGAGAAATATGTTCAAACTGTGAAGTCAATCAAAAGGTGGTCCCGATCAGCGG 180
 QY 305 TGAGCCCTTATCTTACCTTTCAATCCATCCTCGGATGTGCTGCTCCCTTGCATTAAGGCCATA 364
 DB 181 TGAGCCCTTATCTTACCTTTCAATCCATCCTCGGATGTGCTGCTCCCTTGCATTAAGGCCATA 240
 QY 365 TGGTTAAAGGTGTGATGAAGCAACCATTCATTCATTAAGTGAAGCAACATTCATTAAGCAACATTCAC 424
 DB 241 TGGTTAAAGGTGTGATGAAGCAACCATTCATTCATTAAGTGAAGCAACATTCATTAAGCAACATTCAC 300
 QY 425 AGCGTCAACAGATCAAAAGCAGCATATCTCCAGAAACAGGAAAGCCCTGGATGAACAC 484
 DB 301 AGCGTCAACAGATCAAAAGCAGCATATCTCCAGAAACAGGAAAGCCCTGGATGAACAC 360
 QY 485 TGAAGAAACCCCTTACAGGTACCTTGAAGAGGTGTTTGTAGCTCTGCTTAAACATCCAG 544
 DB 361 TTAAGAAACCCCTTACAGGTACCTTGAAGAGGTGTTTGTAGCTCTGCTTAAACATCCAG 420
 QY 545 CGCAATTTGATGCTGATGAATCTGCTGCTCCATGAAGGGCCTTGGAACTGATGAAGATA 604
 DB 421 CGCAATTTGATGCTGATGAATCTGCTGCTCCATGAAGGGCCTTGGAACTGATGAAGATA 480
 QY 605 CTCATATTGAGATTTGGCATCAAGAACTAAACAAAGAAATCAGAGACATTAACAGGCTCT 564
 DB 481 CTCATATTGAGATTTGGCATCAAGAACTAAACAAAGAAATCAGAGACATTAACAGGCTCT 540
 QY 665 ACAGAGAGAACTGAAGAGAGATCTGGCCAAAGACATAACCTCAGACACATCTGGAGATT 724
 DB 541 ACAGAGAGAACTGAAGAGAGATCTGGCCAAAGACATAACCTCAGACACATCTGGAGATT 600
 QY 725 TTCGGAACGCTTTGCTTTCTTTGCTAAGGGTGAACCGATCTCAGAGACTTTGGTGAATG 784
 DB 601 TTCGGAACGCTTTGCTTTCTTTGCTAAGGGTGAACCGATCTCAGAGACTTTGGTGAATG 660
 QY 785 AAGACTTGGCTGATTCAGATGCCAGGGCCTTGTATGAAGCAGAGAAAGGAGAAAGGGA 844
 DB 661 AAGACTTGGCTGATTCAGATGCCAGGGCCTTGTATGAAGCAGAGAAAGGAGAAAGGGA 720
 QY 845 CAGACGTAAACGTTGTTCAATACCATCTTACCACCAAGAGCTATCCCAACTTCGCAGAG 904
 DB 721 CAGACGTAAACGTTGTTCAATACCATCTTACCACCAAGAGCTATCCCAACTTCGCAGAG 780
 QY 905 TGTTTCAGAAATACACCAAGTACAGTAAAGCAATGAACAAAGTTCTGGACCTGGAGT 964
 DB 781 TGTTTCAGAAATACACCAAGTACAGTAAAGCAATGAACAAAGTTCTGGACCTGGAGT 840
 QY 965 TGAAGGTGACATTCAGAAATGCTTCACAGCTATCTGGAAGTGGCCCAACAAAGCAACCCAG 1024
 DB 841 TGAAGGTGACATTCAGAAATGCTTCACAGCTATCTGGAAGTGGCCCAACAAAGCAACCCAG 900
 QY 1025 CTTTCTTTGAGAGAGAGCTTCATCAAGCCATGAAGGTGTTGGAACTGCGCATAGGCAT 1084
 DB 901 CTTTCTTTGAGAGAGAGCTTCATCAAGCCATGAAGGTGTTGGAACTGCGCATAGGCAT 960
 QY 1085 TGATCAGGATATGTTTCCCGTTCGAAATGACATGAATGATATCAAAAGCAATTTCTATC 1144
 DB 961 TGATCAGGATATGTTTCCCGTTCGAAATGACATGAATGATATCAAAAGCAATTTCTATC 1020
 QY 1145 AGAAGATGATGATATCTCCCTTTGCGAAGCCATCCTGGATGAACCAAGAGAGATATG 1204
 DB 1021 AGAAGATGATGATATCTCCCTTTGCGAAGCCATCCTGGATGAACCAAGAGAGATATG 1080
 QY 1205 AGAAATCTGCTGCTCTTTGCGAAGAACTAAACATTCCTTGTGTTCTCAAGCTA 1264
 DB 1081 AGAAATCTGCTGCTCTTTGCGAAGAACTAAACATTCCTTGTGTTCTCAAGCTA 1140

```

QY 1265 TGATCAGAGACGCTTTAATATATATTTTTCATCTCTATAGCTTAAATAGGAAAGTTTCTTC 1324
DB 1141 TGATCAGAGACGCTTTAATATATATTTTTCATCTCTATAGCTTAAATAGGAAAGTTTCTTC 1200
QY 1325 AACAGGATACAGGTGATGCTACCTGCTGAAATATAGCTTTAAATCAATTTTAT 1384
DB 1201 AACAGGATACAGGTGATGCTACCTGCTGAAATATAGCTTTAAATCAATTTTAT 1260
QY 1385 ATTATAACTCTGTATATAGAGATAGTCCATTTTAAATATGTTTTTCCCAAAACCAT 1444
DB 1261 ATTATAACTCTGTATATAGAGATAGTCCATTTTAAATATGTTTTTCCCAAAACCAT 1320
QY 1445 AAACCCCTATCAAGTTGTTCTAGTAACTATCAATGAGAAAGATGCTATGATGCTGAAA 1504
DB 1321 AAACCCCTATCAAGTTGTTCTAGTAACTATCAATGAGAAAGATGCTATGATGCTGAAA 1380
QY 1505 TAAATGNCGTC 1516
DB 1381 TAAATGACGTC 1392

RESULT 8
US-10-133-937-2
; Sequence 2, Application US/10133937
; Publication No. US20030207278A1
; GENERAL INFORMATION:
; APPLICANT: Khan, Javed
; APPLICANT: Ringner, Markus
; APPLICANT: Peterson, Carsten
; APPLICANT: Meltzer, Paul
; TITLE OF INVENTION: METHODS FOR ANALYZING HIGH DIMENSIONAL DATA FOR CLASSIFYING,
; TITLE OF INVENTION: DIAGNOSIS, PROGNOSTICATING, AND/OR PREDICTING DISEASES AND
; TITLE OF INVENTION: OTHER BIOLOGICAL STATES
; FILE REFERENCE: 11613.5US01
; CURRENT APPLICATION NUMBER: US/10/133,937
; CURRENT FILING DATE: 2002-11-04
; NUMBER OF SEQ ID NOS: 99
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1399
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-133-937-2

Query Match 91.5%; Score 1387.8; DB 15; Length 1399;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1389; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 125 AGTGTGAATCTTCAGAGAGAAATTTCTCTTTAGTTCTTTGCAAGAGGTAGAGATAAG 184
DB 1 AGTGTGAATCTTCAGAGAGAAATTTCTCTTTAGTTCTTTGCAAGAGGTAGAGATAAG 60
QY 185 ACATTTTTCAAAATGCGCAATGGTATCAGAAATTCCTCAAGCAGGCGCTGGTTATTGAAA 244
DB 61 ACATTTTTCAAAATGCGCAATGGTATCAGAAATTCCTCAAGCAGGCGCTGGTTATTGAAA 120
QY 245 ATGAAGAGCGAGGAATATGTTTCAAAGTGTGAAGTATCCTCAAGAGGTGTCGGATCAGCGG 304
DB 121 ATGAAGAGCGAGGAATATGTTTCAAAGTGTGAAGTATCCTCAAGAGGTGTCGGATCAGCGG 180
QY 305 TGAGCCCTATCTTACCTTCAATCCATCCCTCGATGTCGCTGCTTGCATAGGCCATAA 364
DB 181 TGAGCCCTATCTTACCTTCAATCCATCCCTCGATGTCGCTGCTTGCATAGGCCATAA 240
QY 365 TGGTTAAAGTGTGGATGAAGCAACCATCATTTGACATTTAACTAAGCGAACAATGCAAC 424
DB 241 TGGTTAAAGTGTGGATGAAGCAACCATCATTTGACATTTAACTAAGCGAACAATGCAAC 300
QY 425 AGCTGCAACAGATCAAGAGAGCATATCTCCAGGAAACAGGAAAGCCCTGGATGAAACAC 484
DB 301 AGCTGCAACAGATCAAGAGAGCATATCTCCAGGAAACAGGAAAGCCCTGGATGAAACAC 360

```

```

QY 485 TGAAGAAAGCCCTTACAGGTCACTTGAAGAGGTGTTTTAGCTCTGCTAAAAAATCCAG 544
DB 361 TTAAGAAAGCCCTTACAGGTCACTTGAAGAGGTGTTTTAGCTCTGCTAAAAAATCCAG 420
QY 545 CGCAATTTGATGCTGATGAACCTTCTGCTGCCATGAAGGSCCTTGGAACTGATGAAGATA 604
DB 421 CGCAATTTGATGCTGATGAACCTTCTGCTGCCATGAAGGSCCTTGGAACTGATGAAGATA 480
QY 605 CTCTAAATTTGAGATTTTGGCATCAAGAACTAAACAAAGAAATCAGAGACATTTAAACAGGCTCT 664
DB 481 CTCTAAATTTGAGATTTTGGCATCAAGAACTAAACAAAGAAATCAGAGACATTTAAACAGGCTCT 540
QY 665 ACAGAGAGGAACTGAGAGAGATCTGCCAAAGACATTAACCTCAGACACATCTGGAGATT 724
DB 541 ACAGAGAGGAACTGAGAGAGATCTGCCAAAGACATTAACCTCAGACACATCTGGAGATT 600
QY 725 TTGGAACGCTTTGCTTTCTCTTCTTAAGGCTGACCGATCTGAGGACTTTGGTGTGAATG 784
DB 601 TTGGAACGCTTTGCTTTCTCTTCTTAAGGCTGACCGATCTGAGGACTTTGGTGTGAATG 660
QY 785 AAGACTTTGGCTGATTTCAGATGCCAGGSCCTTGTATGAAGCAGGAGAAAGGAGAAAGGGA 844
DB 661 AAGACTTTGGCTGATTTCAGATGCCAGGSCCTTGTATGAAGCAGGAGAAAGGAGAAAGGGA 720
QY 845 CAGACGTAAACGCTGTTCAATACCATCTTACCACCAAGAGCTATCCACAACTTTCGCAGAG 904
DB 721 CAGACGTAAACGCTGTTCAATACCATCTTACCACCAAGAGCTATCCACAACTTTCGCAGAG 780
QY 905 TGTTTCAGAAATACACCAAGTACAGTGAAGTACATGACATGAACAAAGTTCTGCACTGGAGT 964
DB 781 TGTTTCAGAAATACACCAAGTACAGTGAAGTACATGACATGAACAAAGTTCTGCACTGGAGT 840
QY 965 TGAAGAGTGACATTTGAGAAATGCTCAGACGTATCTGTAAGTGGCCGACCAAGCAAAACAG 1024
DB 841 TGAAGAGTGACATTTGAGAAATGCTCAGACGTATCTGTAAGTGGCCGACCAAGCAAAACAG 900
QY 1025 CTTTCTTTGAGAGAACTTCAACAGCCATGAAGGTGTTGAACTCGGCCATAGGCGAT 1084
DB 901 CTTTCTTTGAGAGAACTTCAACAGCCATGAAGGTGTTGAACTCGGCCATAGGCGAT 960
QY 1085 TGATCAGGATTTATGTTTCCCGTTCTGAAATTTGACATGAATGATATCAAAAGCATTTCTATC 1144
DB 961 TGATCAGGATTTATGTTTCCCGTTCTGAAATTTGACATGAATGATATCAAAAGCATTTCTATC 1020
QY 1145 AGAAGATGATGTTATCTCCCTTTGCAAGCCATCTGGATGAACCAAGGAGAGATAG 1204
DB 1021 AGAAGATGATGTTATCTCCCTTTGCAAGCCATCTGGATGAACCAAGGAGAGATAG 1080
QY 1205 AGAAATCCCTGGTGGCTTTTGTGAGGAACTAAACATTTCCCTTGTATGCTCAAGCTA 1264
DB 1081 AGAAATCCCTGGTGGCTTTTGTGAGGAACTAAACATTTCCCTTGTATGCTCAAGCTA 1140
QY 1265 TGATCAGAGACTTTAATATATATTTTTCATCTCTATAGCTTAAATAGGAAAGTTTCTTC 1324
DB 1141 TGATCAGAGACTTTAATATATATTTTTCATCTCTATAGCTTAAATAGGAAAGTTTCTTC 1200
QY 1325 AACAGGATACAGGTGATGCTACCTGCTGAAATATAGCTTTAAATCAATTTTAT 1384
DB 1201 AACAGGATACAGGTGATGCTACCTGCTGAAATATAGCTTTAAATCAATTTTAT 1260
QY 1385 ATTATAACTCTGTATATAGAGATAAGTCCATTTTAAATATGTTTTTCCCAAAACCAT 1444
DB 1261 ATTATAACTCTGTATATAGAGATAAGTCCATTTTAAATATGTTTTTCCCAAAACCAT 1320
QY 1445 AAACCCCTATCAAGTTGTTCTAGTAACTATCAATGAGAAAGATGCTATGATGCTGAAA 1504
DB 1321 AAACCCCTATCAAGTTGTTCTAGTAACTATCAATGAGAAAGATGCTATGATGCTGAAA 1380
QY 1505 TAAATGNCGTC 1516
DB 1381 TAAATGACGTC 1392

```


[illegible]

Db 1321 GTAAACAATCATGAGAAAGATGCTTATGTAGCTGAAATTAATTAATGAGCTC 1370

RESULT 12

US-09-917-800A-1585

; Sequence 1585, Application US/09917800A

; Patent No. US20020119462A1

; GENERAL INFORMATION:

; APPLICANT: Mendrick, Donna

; APPLICANT: Porter, Mark

; APPLICANT: Johnson, Kory

; APPLICANT: Castle, Arthur

; APPLICANT: Elashoff, Michael

; APPLICANT: Gate Logic, Inc.

; TITLE OF INVENTION: Molecular Toxicology Modeling

; FILE REFERENCE: 44921-5038-US

; CURRENT APPLICATION NUMBER: US/09/917,800A

; CURRENT FILING DATE: 2001-07-31

; PRIOR APPLICATION NUMBER: US 60/222,040

; PRIOR FILING DATE: 2000-07-31

; PRIOR APPLICATION NUMBER: US 60/222,880

; PRIOR FILING DATE: 2000-11-02

; PRIOR APPLICATION NUMBER: US 60/290,029

; PRIOR FILING DATE: 2001-05-11

; PRIOR APPLICATION NUMBER: US 60/290,645

; PRIOR FILING DATE: 2001-05-15

; PRIOR APPLICATION NUMBER: US 60/292,336

; PRIOR FILING DATE: 2001-05-22

; PRIOR APPLICATION NUMBER: US 60/295,798

; PRIOR FILING DATE: 2001-06-06

; PRIOR APPLICATION NUMBER: US 60/297,457

; PRIOR FILING DATE: 2001-06-13

; PRIOR APPLICATION NUMBER: US 60/298,884

; PRIOR FILING DATE: 2001-06-19

; PRIOR APPLICATION NUMBER: US 60/303,459

; PRIOR FILING DATE: 2001-07-09

; NUMBER OF SEQ ID NOS: 1740

; SOFTWARE: Patentin Ver. 2.1

; SEQ ID NO 1585

; LENGTH: 1402

; TYPE: DNA

; ORGANISM: Rattus norvegicus

; FEATURE:

; OTHER INFORMATION: Genbank Accession No. US20020119462A1 NM_012904

US-09-917-800A-1585

Query Match 58.5%; Score 886.8; DB 9; Length 1402;

Best Local Similarity 80.1%; Pred. No. 3.9e-230;

Matches 1101; Conservative 0; Mismatches 267; Indels 6; Gaps 5

Qy 140 GAGAAGAAATTTCTCTTTAGTTCTTTGCAAGA-AGGTAGAGNTAAAGACACTTTTCAAAA 198

Db 2 GAGCAAAAGCTTCTCTTCAGTTCCCTGGAAGCAAGGCAATACAAGATACATTTTATTTAAAA 61

Qy 199 ATGCAATGGTATCAGAAATTCCTCAAGCAGGCCCTGGTTTATTGAAATGAAGAGCAGGAA 258

Db 62 ATGCAATGGTATCAGAAATTCCTCAAGCAGGCCCTGCTATATTGAAAGCAAGAGCAGGAA 121

Qy 259 TATGTTCAAACTGTGAAAGTCAATCAAAAGGTGGTCCCGGATCAGCGGTGAGGCCCTATCCT 318

Db 122 TATGTTCAAGCTGTAAAAATCCTCAAAAGGTGGTCTCTGGATCAGCAGTGAAGCCCTACCT 181

Qy 319 ACCTTCAATCCATCCTCGGATGTCGCTGCCCTTGCATAAGGCCATAATGGTTAAAGGTGTG 378

Db 182 TCCTTCAATCCGTCCTCGATGTGTGCTGCTTGCATCAAGCTATCAUGTTAAAGTGTG 241

Qy 379 GATGAGCAACCATCATGTGACATTTCTAACTAAGCGAAACAATGACAGCGTCAACAGATC 438

Db 242 GATGAGGCAACCATCATGTGACATCCTTACCAAGAGAACCAATGTCTCAGCGCCAGCAGATC 301

Qy 439 AAAGCAGCATATCTCCAGGAACAGGAAGCCCTGGATGAAACACTCAAGAAAGCCCTT 498

Db 302 AAGGAGCATCTTACAGGAGACTGGGAGCCCTGGATGAACCTTGAAAAAGCCCTT 361
 QY 499 ACAGGTCACTTACAGGAGTGTGTTTAGCTCTGCTAAATAATCCAGCGCAATTTGATCCT 558
 Db 362 ACGGCGCACTGGAGAGTGTGTTTGGCTATGCTCAAGACCCAGCTCAGTTGATGCA 421
 QY 559 GATGAATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 618
 Db 422 GATGAATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 481
 QY 619 TTGGCATCAAGAACTAAACAAAGAAATCAGAGACATTAACAGGGTCTACAGAGAGAACTG 678
 Db 482 TTGACAAACAGATCTAACCCAGCAATCAGAGAGATTAAGAGTCTACAGAGAGAGCTG 541
 QY 679 AAGAGAGTCTGCGCAAGACATAAATCAGAGACATCTGAGAGATTTTCGGAACCTTTG 738
 Db 542 AAAAGAGATCTGCGCAAGACATCACTTCGACACATCTGGAGACTTTTCGTAATGCTT 601
 QY 739 CTTTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 798
 Db 602 CTTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 661
 QY 799 TCAGATGCCAGGCTTGTATGAAGCAGAGAAAGGAGAAAGGAGGACAGACGTAACGCTG 858
 Db 662 ACAGATGCCAGGCTTGTATGAAGCAGAGAAAGGAGAAAGGAGGACAGACGTAACGCTG 721
 QY 859 TTTCAATACCATCTTACCAACCAAGCTATCCACACTTTCGAGAGTGTTCAGAAATAC 918
 Db 722 TTTCAATACCATCTTACCAACCAAGCTATCCACACTTTCGAGAGTGTTCAGAAATAC 781
 QY 919 ACCAAGTACAGTGAAGCATGACATGAACAAAGTTCGACCTCGAGTGTGAAGGTGACAT 978
 Db 782 AGAAGTACAGTGAACATGACATGAACAAAGCTGCTGATCTGGAATGAAGGTGACAT 841
 QY 979 GAGAAATGCTCTACAGTATCTGAGTGAAGTGGCCACAGCAACCAAGCTTCTTCGAG 1038
 Db 842 GAGAAGTGCCTCAACCAATGTTGAAGTGTGGCCACAGCAACCAAGCTTCTTCGAG 901
 QY 1039 AAGCTTTCATCAAGCCATGAAGGTGTTGGAATCGCCATGAAGCAATGATCAGGATATG 1098
 Db 902 AACTGTATGAAGCCATGAAGGTGTTGGAATCGCCATGAAGCAATGATCAGGATATG 961
 QY 1099 GTTTCGGTCTGAAATGACATGATGATCAAGCAATCTATCAGAGATGATGCT 1158
 Db 962 GTCTCCGTTTCGAAATGACATGATGATCAAGCAATCTATCAGAGATGATGCT 1021
 QY 1159 ATCTCCCTTTGCCAAGCATCTCGGATGAACCAAGGAGAGATGATCAGAAATCTCTG 1218
 Db 1022 ATCTCCCTTTGCCAAGCATCTCGGATGAACCAAGGAGAGATGATCAGAAATCTCTG 1081
 QY 1219 GCTCTTTGCGAGAAACTAAACATCTCTGATGATGATCAAGCTATGATCAGAGAA-CT 1277
 Db 1082 GCTCTGTTGCGAGAAACTAAACATCTCTGATGATGATCAAGCTATGATCAGAGAA-CT 1141
 QY 1278 TTAATATATATTTTCACTCTATGATGATGATGATGATGATGATGATGATGATGATGAT 1337
 Db 1142 CTTAGCGTGTGTTTCTCTATGATGATGATGATGATGATGATGATGATGATGATGAT 1201
 QY 1338 TGTAGTACCTACATGCTGAGAAATATAGCTTTTAAATCAATTTTATATATATATATAT 1397
 Db 1202 TCTAATACCTTCTT--TGAATATGATGATGATGATGATGATGATGATGATGATGATGAT 1259
 QY 1398 ATATATGAGATA-AGTCCATTTTAAATATGTTTTCCTCAACCAAGCTATGATGATGATGAT 1455
 Db 1260 ACAAATGAGAAATATCTGTTTGTATGATGATGATGATGATGATGATGATGATGATGAT 1319
 QY 1456 AAGTGTCTAGTAAATACATGAGAAAGATGCTATGATGATGATGATGATGATGATGATGAT 1509
 Db 1320 CAAGTCACTTTGATCACTTCTCGAGAAAGAGTTTACATAGATAAATAAATAA 1373

; Sequence 299, Application US/10316253
 ; Publication No. US20030162706A1
 ; GENERAL INFORMATION:
 ; APPLICANT: The Procter & Gamble Company
 ; APPLICANT: Peters, Kevin
 ; APPLICANT: Thompson, Larry
 ; APPLICANT: Wang, Feng
 ; APPLICANT: Greis, Kenneth
 ; TITLE OF INVENTION: Angiogenesis Modulating Proteins
 ; FILE REFERENCE: 8865M
 ; CURRENT APPLICATION NUMBER: US/10/316,253
 ; CURRENT FILING DATE: 2002-12-10
 ; PRIOR APPLICATION NUMBER: US 60/355,295
 ; PRIOR FILING DATE: 2002-02-08
 ; NUMBER OF SEQ ID NOS: 308
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 299
 ; LENGTH: 1402
 ; TYPE: DNA
 ; ORGANISM: Rattus norvegicus
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (62)..(1102)
 ; OTHER INFORMATION:
 ; US-10-316-253-299

 Query Match 58.5%; Score 886.8; DB 14; Length 1402;
 Best Local Similarity 80.1%; Pred. No. 3.9e-230;
 Matches 1101; Conservative 0; Mismatches 267; Indels 6; Gaps 5;

 QY 140 GAGAAGAAATTCCTTTAGTCTTTCGAGA-AGGTAGATGAAGACATTTTCAAAA 198
 Db 2 GAGCAAGCTTCTCTTCAGTTCCTGGAAGACAGGCAATCAAGATATCTTTTAAAA 61
 QY 199 ATGGCAATGTATCAGAAATTCCTCAAGCAGGCTGTGTTTATTGAAAATGAAGACAGGAA 258
 Db 62 ATGGCAATGTATCAGAAATTCCTCAAGCAGGCTGTGTTTATTGAAAATGAAGACAGGAA 121
 QY 259 TATGTTCAAACTGTGAAGTCACTCAAAAGTGTTCGCGATCAGGGTGAAGGCTTATCCT 318
 Db 122 TATGTTCAAGCTGTAAAAATCCTCAAAAGTGTTCGCGATCAGGGTGAAGGCTTATCCT 181
 QY 319 ACCTTCAATCCTCCTCGATGTGCTGCTTGCATGAAGGCCATTAATGTTAAAGGTGTG 378
 Db 182 TCCTTCAATCGTCTCGATGTGCTGCTTGCATGAAGGCCATTAATGTTAAAGGTGTG 241
 QY 379 GATGAAGCAACCATCATCTTAATCTTAAGTGAAGGAAACATGCAAGCTCAAGATC 438
 Db 242 GATGAGCAACCATCATCTTAATCTTAAGTGAAGGAAACATGCAAGCTCAAGATC 301
 QY 439 AAAGCAGCATATCTCCAGGAAACAGGAAAGCCCTGATGAACACCTGAAGAAAGCCCTT 498
 Db 302 AAGCAGCATATCTTAAGGAGACTGGAAGGCCCTGATGAAGCCCTGAAGAAAGCCCTT 361
 QY 499 ACAGTCACTTGAAGGAGTGTGTTTATGCTTGTCTGCTTAAAACTCAGCGCAATTTGATGCT 558
 Db 362 ACGGCGCACTTGAAGGAGTGTGTTTATGCTTGTCTGCTTCAAGACCCAGCTCAGTTGATGCA 421
 QY 559 GATGAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 618
 Db 422 GATGAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 481
 QY 619 TTGGCATCAAGAACTAAACAAAGAAATCAGAGACATTAACAGGGTCTACAGAGAGAACTG 678
 Db 482 TTGACAAACAGATCTAACCCAGCAATCAGAGAGATTAAGAGTCTACAGAGAGAGCTG 541
 QY 679 AAGAGAGTCTGCGCAAGACATAAATCAGAGACATCTGAGAGATTTTCGGAACGCTT 738
 Db 542 AAAAGAGATCTGCGCAAGACATCACTTCGACACATCTGGAGACTTTTCGTAATGCTT 601
 QY 739 CTTTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 798
 Db 602 CTTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 661

301	DB	GAATACTCTGGTGGCTCTTTGTGGAGAACTTAAACATTCCTTGATGTCTCAAGCTAT	360
1265	QY	GATCAGAGACATTTAAATATATATATTTTCATCTATATAGCTTAAATAGGAAAGTTTCTTCA	1325
361	DB	GATCAGAAGACATTTAAATATATATATTTTCATCTATATAGCTTAAATAGGAAAGTTTCTTCA	420
1326	QY	ACAGGATTAACAGTGTAGCTACCTACATGCTGAAAAATATAGCCTTTAAATCATTTTATA	1385
421	DB	ACAGGATTAACAGTGTAGCTACCTACATGCTGAAAAATATAGCCTTTAAATCATTTTATA	480
1386	QY	TTATAACTCTGTATTAATACAGATAAGCTCCATTTTTTAAAAATGTTTTCCCAACCATAA	1445
481	DB	TTATAACTCTGTATTAATACAGATAAGCTCCATTTTTTAAAAATGTTTTCCCAACCATAA	540
1446	QY	AACCCCTATAAAGTTGTTCTTAGTAAACAATACATGAGAAAGATGCTCTATGTAGCTGAAAT	1505
541	DB	AACCCCTATAAAGTTGTTCTTAGTAAACAATACATGAGAAAGATGCTCTATGTAGCTGAAAT	600

Search completed: March 12, 2004, 06:13:12
Job time : 358.368 secs

Search completed: March 12, 2004, 06:13:12
Job time : 358.368 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

M nucleic - nucleic search, using sw model

run on: March 10, 2004, 09:18:38 ; Search time 2509.9 Seconds

(without alignments)

18036.975 Million cell updates/sec

Title: US-10-084-817-8

Perfect score: 1516

Sequence: 1 cttgttttggacatagct.....gctgaaataaaatgncgtc 1516

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_estc:*

9: gb_estl:*

10: gb_est2:*

11: gb_est3:*

12: gb_est4:*

13: gb_est5:*

14: gb_est6:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gssl:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	1030.6	68.0	1201	9	AL570884
C 2	1019	67.2	1169	9	AL570428
C 3	1014.4	66.9	1201	13	BX438944
4	1009.6	66.6	1128	13	BUS08007

5	1002	66.1	1183	9	AL544231
6	988.6	65.2	1163	13	BU902298
7	985.8	65.1	1201	9	AL553114
C 8	985.8	65.1	1201	9	AL575425
9	984	64.9	1201	9	AL541874
C 10	971.4	64.3	1136	9	AL574991
11	959.8	63.3	1201	9	AL553129
12	958.2	63.2	1201	9	AL544959
13	956	63.1	1164	9	AL551427
C 14	955.8	63.0	1201	9	AL576223
C 15	946.6	62.4	1201	9	AL553095
16	942	62.1	1076	12	BM476344
17	940.8	62.1	1077	12	BM473319
18	937.6	61.8	1201	9	AL551980
19	930.2	61.4	996	13	BQ222672
20	926.6	61.1	987	9	AL540651
C 21	925.2	61.0	1201	9	AL568911
22	910	60.0	1085	12	BM550642
23	906.4	59.8	1115	14	CK232619
24	903.6	59.6	990	13	BQ228162
C 25	901.4	59.5	1201	9	AL541873
26	897.4	59.2	1201	9	AL541443
27	894.2	59.0	1052	12	BM803761
28	893.8	59.0	1401	11	AK002681
29	891.4	58.8	1044	12	BM803483
30	890	58.0	939	12	BM468858
31	879.2	58.0	1059	14	CK232521
32	878.8	58.0	1201	9	AL542692
33	875.8	57.8	1186	9	AL553164
34	874.2	57.7	1033	12	BM460109
C 35	867.8	57.2	1181	9	AL576340
C 36	867.6	57.2	917	9	AL576603
C 37	866.4	57.2	973	9	AL568522
38	864.2	57.0	935	13	BQ959907
39	863	56.9	894	13	BU173401
40	857.8	56.6	880	13	BU902570
41	855.8	56.5	1054	12	BM458053
42	846.8	55.9	911	12	BM461778
43	840.4	55.4	942	9	AL553679
44	836	55.1	1044	12	BM558842
45	833	54.9	909	13	BU190139

ALIGNMENTS

RESULT 1
AL570884/c
LOCUS
DEFINITION
AL570884 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA clone CS0DI012YP05 3-PRIME, mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1201 bp mRNA linear EST 31-MAY-2003

AL570884 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA clone CS0DI012YP05 3-PRIME, mRNA sequence.

AL570884
EST.
Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 16, 2001 this sequence version replaced gi:12927628.

Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 5436.f For more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DI012CH03NP1&cluster=5436.f. Contact :
Peng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0D1012CH03NP1.

FEATURES

source
1. .1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1012YP05"
/tissue type="PLACENTA COT 25-NORMALIZED"
/clone lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 68.0%; Score 1030.6; DB 9; Length 1201;
Best Local Similarity 93.2%; Pred. No. 1.2e-263;
Matches 1044; Conservative 33; Mismatches 41; Indels 2; Gaps 1;
QY 387 AACCATCTGACATTCCTAACTAAGCGAACAATGACAGCGTCAACAGATCAAGCAGC 446
DB 1122 AASCACCATATTACATTTTAMTWAGSAAACAATCACCGSKTCAMVAGATWAAGMCAGC 1063
QY 447 ATATCTCCAGGAACAGGAAGCCCTGGATGAACACTGAAGAAAGCCCTTACAGGTCA 506
DB 1062 ATTTCCTCGAAGCAAGAAACCCCTGGTGAACACTGAAGAAACCCCTTACAGGTCA 1003
QY 507 CTTGAGGAGGTGTTTAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 566
DB 1002 CTTTGGGATGTTTCTTTAGVCTTSCWAAACACCCAGCGCAATTTGATGCTGATGA 943
QY 567 TCGTGTCCATGAGCGGCTTGGAGCTGATGAGTACTCTAATTCAGATTTTGGCATC 526
DB 942 TCGTGCTSCATGAGGCGCTTGGAGCTGATGAGTACTCTAATTCAGATTTTGGCATC 583
QY 627 AAGAACCTAACAAAGAAATCAGAGACATTAAACAGGCTCTACAGAGAGGAATCGAAGAGA 686
DB 882 AAGAACCTAACAAAGAAATCAGAGACATTAAACAGGCTCTACAGAGAGGAATCGAAGAGA 823
QY 687 CTGGCCAAAGACATACCTCAGACATCTGGAGATTTTCGGAAGCCTTGGCTTCTCT 746
DB 822 TCTGCGCAAGACATACCTCAGACATCTGGAGATTTTCGGAAGCCTTGGCTTCTCT 763
QY 747 TGCTAAGGCTGACCGATCTGAGGACTTTGGTGTGAATGAAGACTTTGGCTGATTCAGATGC 806
DB 762 TGCTAAGGCTGACCGATCTGAGGACTTTGGTGTGAATGAAGACTTTGGCTGATTCAGATGC 703
QY 807 CAGGCGCTTGATGAGCAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 866
DB 702 CAGGCGCTTGATGAGCAG 643
QY 867 CATCTTACCACCAAGAGCTATCCACAACTTCGAGAGTGTTCAGAAATACACCAAGTA 926
DB 642 CATCTTACCACCAAGAGCTATCCACAACTTCGAGAGTGTTCAGAAATACACCAAGTA 583
QY 927 CAGTAAGCATGACATGAACAAAGTTCTGGACCTGGAGTTGAAGGTGACATTCAGAAATG 986
DB 582 CAGTAAGCATGACATGAACAAAGTTCTGGACCTGGAGTTGAAGGTGACATTCAGAAATG 523
QY 987 CCTCAGAGTATCGTGAAGTGGCGCAGACAAACAGCTTTCTTTCAGAGAGGCTTCA 1046
DB 522 CCTCAGAGTATCGTGAAGTGGCGCAGACAAACAGCTTTCTTTCAGAGAGGCTTCA 463
QY 1047 TCAAGCCATGAAGGTGTGGAATTCGCCATPAAGGCATTCATCAGGATTCAGTTCCCG 1106
DB 462 TCAAGCCATGAAGGTGTGGAATTCGCCATPAAGGCATTCATCAGGATTCAGTTCCCG 403
QY 1107 TCTGAAATTCAGATGATATCAAGCATTCATCAGAGATGATGATGATGATGATGATGATG 1166
DB 402 TCTGAAATTCAGATGATATCAAGCATTCATCAGAGATGATGATGATGATGATGATGATG 343
QY 1167 TTGCAAGCCATCTCGGATGAACAAAGGAGAGATGATGAGAAATCTCGTGGCTCTTTG 1226

DB 342 TTGCCAAGCCATCTCGGATCAAAACCAAGAGATTATGAGAAATCTCGTGGCTCTTTG 283
QY 1227 TGGAGGAAACTAAACATTCCCTTGTGCTCAAGCTATGATCAGAGACTTTAAATTATA 1286
DB 282 TGGAGGAAACTAAACATTCCCTTGTGCTCAAGCTATGATCAGAGACTTTAAATTATA 223
QY 1287 TATTTTCATCTATAAGCTTAAATAGGAAAGTTCTTCAACAGAGTTACAGTGTAGCTAC 1346
DB 222 TATTTTCATCTATAAGCTTAAATAGGAAAGTTCTTCAACAGAGTTACAGTGTAGCTAC 163
QY 1347 CTACATCTGAAATAATAGCTTTAAATCATTTTATATATATAAATCTCTGTATAAGAG 1406
DB 162 CTACATCTGAAATAATAGCTTTAAATCATTTTATATATATAAATCTCTGTATAAGAG 103
QY 1407 ATAAGTCCATTTTAAAAATGTTTTCCCAACCATATAAAACCTATACAGTTGTTCTA 1466
DB 102 ATAAGTCCATTTTAAAAATGTTTTCCCAACCATATAAAACCTATACAGTTGTTNTN 43
QY 1467 G--TAACAATACATGAGAAAGATGCTATGATAGCTGAAA 1504
DB 42 AGTNACAACATGAGAAAGATGATATATGTDNNDNND 3
RESULT 2
AL570428/c
AL570428/c
LOCUS
DEFINITION
AL570428 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1019Yf12 3-PRIME, mRNA sequence.
AL570428
AL570428.2 GI:31291848
ACCESSION
VERSION
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1169)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
COMMENT
On Feb 16, 2001 this sequence version replaced gi:12926726.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 5436.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0D1019DC06NP1&cluster=5436.f. Contact :
Feng Liang Email : fliang@lifetech.com URL : Corporation 1600
http://fulllength.invitrogen.com/ Invitrogen
Faraday Avenue Genoscope sequence ID : CS0D1019DC06NP1.
Location/Qualifiers
1. .1169
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1019Yf12"
/tissue type="PLACENTA COT 25-NORMALIZED"
/clone lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
FEATURES
source
Query Match 67.2%; Score 1019; DB 9; Length 1169;
Best Local Similarity 98.3%; Pred. No. 1.5e-260;
Matches 1044; Conservative 5; Mismatches 11; Indels 2; Gaps 2;
QY 434 AGATCAAGCAGCATATCTCCAGGAACAGAAAGCCCTCGATGACACTGAGAAG 493

562 GAACTTCGTCTGCCATGAAGGCGCTTGGAACTGATGAAGATACTCTAATTGAGATTTTG 621
 Db GAACTTCGTCTGCCATGAAGGCGCTTGGAACTGATGAAGATACTCTAATTGAGATTTTG 543
 QY GATCAGAACTAACAAAGAAATCAGAGACATTAACAGGGTCTACAGAGAGAACTGAAG 581
 Db GATCAGAACTAACAAAGAAATCAGAGACATTAACAGGGTCTACAGAGAGAACTGAAG 503
 QY AGAGATCTGCCAAAGACATAAAGCTCAGACACATCTGGAGATTTTCGGAACGCTTTGCTT 741
 Db AGAGATCTGCCAAAGACATAAAGCTCAGACACATCTGGAGATTTTCGGAACGCTTTGCTT 663
 QY TCTCTTCTTAAGGGTACCGATCTGAGGACTTTGGTGTGATGAAGACTTGGCTGATTC 801
 Db TCTCTTCTTAAGGGTACCGATCTGAGGACTTTGGTGTGATGAAGACTTGGCTGATTC 723
 QY GATGCCAGGCGCTTGTATGAAGCAGAGAGAAAGGAGAGAGGAGACAGCTAAACGCTGTC 861
 Db GATGCCAGGCGCTTGTATGAAGCAGAGAGAAAGGAGAGAGGAGACAGCTAAACGCTGTC 783
 QY AATACCATCTTACCACAGAGAGCTATCCAACTTCGAGAGTGTTCAGAAATACACC 921
 Db AATACCATCTTACCACAGAGAGCTATCCAACTTCGAGAGTGTTCAGAAATACACC 843
 QY AAGTACAGTAAAGATGACATGAACAAAGTCTCGACCTGGAGTTGAAAGTGACATTTGAG 981
 Db AAGTACAGTAAAGATGACATGAACAAAGTCTCGACCTGGAGTTGAAAGTGACATTTGAG 903
 QY AATGCTCTCAGATATCGTGAAGTGGCCCAAGCAAAACAGCTTCTTTGAGAGAG 1041
 Db AATGCTCTCAGATATCGTGAAGTGGCCCAAGCAAAACAGCTTCTTTGAGAGAG 963
 QY CTTCATCAAGCCATGAAGGTTTGGAACTCGCATGAAGCAATTCATCAGGATTTGTT 1101
 Db CTTCATCAAGCCATGAAGGTTTGGAACTCGCATGAAGCAATTCATCAGGATTTGTT 1023
 QY TCCGTTTCTGAAATGACATGAATATATCAATCAATTTATCAGAAAGATGTGKDTYT 1083
 Db TCCGTTTCTGAAATGACATGAATATATCAATCAATTTATCAGAAAGATGTGKDTYT 1083
 QY TCCGTTTCTGAAATGACATGAATATATCAATCAATTTATCAGAAAGATGTGKDTYT 1221
 Db TCCGTTTCTGAAATGACATGAATATATCAATCAATTTATCAGAAAGATGTGKDTYT 1139
 QY TTTGTTGAGGAGAA 1235
 Db TTTGTTGAGGAGAA 1153

RESULT 4
 BU508007
 LOCUS
 DEFINITION AGENCOURT 10128424 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6502033
 5', mRNA sequence.
 ACCESSION BU508007
 VERSION BU508007.1 GI:22814240
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1128)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov
 Plate: LLAM14057 Row: h Column: 02
 High quality sequence stop: 828.
 FEATURES
 Location/Qualifiers
 1..1128
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clones="IMAGE:6502033"
 /tissue_type="leiomyosarcoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_71"
 /note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 2.1 Kb."

ORIGIN

Query Match 66.6%; Score 1009.6; DB 13; Length 1128;
 Best Local Similarity 97.8%; Pred. No. 4.9e-258;
 Matches 1076; Conservative 0; Mismatches 19; Indels 5; Gaps 5;
 QY 145 GAATTTCTCTTAGTCTTTTTCGAAAGAGGTAGAGATAAGACA-CTTTTCAAAAATGGC 203
 Db 1 GAATTTCTCTTAGTCTTTTTCGAAAGAGGTAGAGATAAGACANCTTTTCAAAAATGGC 60
 QY 204 AATGGTATCAGAAATTCCTCAAGCAGGCGCTGTTTATTGAAAATGAAGAGCAGGAATATGT 263
 Db 61 AATGGTATCAGAAATTCCTCAAGCAGGCGCTGTTTATTGAAAATGAAGAGCAGGAATATGT 120
 QY 264 TCAAACTGTGAAGTCATCCAAAGAGTGTCCGGATCAGCGGTGAGCCCTATCTACCTT 323
 Db 121 TCAAACTGTGAAGTCATCCAAAGAGTGTCCGGATCAGCGGTGAGCCCTATCTACCTT 180
 QY 324 CAATCCATCTCGGATGTCGCTGCTTGCATAGGCCATAATGGTTAAAGGTGCGATGA 383
 Db 181 CAATCCATCTCGGATGTCGCTGCTTGCATAGGCCATAATGGTTAAAGGTGCGATGA 240
 QY 384 AGCAACCATCATTCGATTTCTAAGCGAAACAATGCAAGCGTCAACAGATCAAGC 443
 Db 241 AGCAACCATCATTCGATTTCTAAGCGAAACAATGCAAGCGTCAACAGATCAAGC 300
 QY 444 AGCATATCTCCAGAAACACAGAAAGCCCTGGATGAACACACTGAAGAAAGCCCTTACAGG 503
 Db 301 AGCATATCTCCAGAAACACAGAAAGCCCTGGATGAACACACTGAAGAAAGCCCTTACAGG 360
 QY 504 TCACCTTGAGGAGGTGTTTGTAGCTCTGTAAATACTCCAGCGCAATTTGATGCTGATGA 563
 Db 361 TCACCTTGAGGAGGTGTTTGTAGCTCTGTAAATACTCCAGCGCAATTTGATGCTGATGA 420
 QY 564 ACTTCGTGCTCCATGAAGGCGCTTGGAACTGATGAAGATACTCTAATTGAGATTTTGGC 623
 Db 421 ACTTCGTGCTCCATGAAGGCGCTTGGAACTGATGAAGATACTCTAATTGAGATTTTGGC 480
 QY 624 ATCAAGAACTAACAAAGAAATCAGAGACATTTAAACAGGGTCTACAGAGAGAACTGAAGAG 683
 Db 481 ATCAAGAACTAACAAAGAAATCAGAGACATTTAAACAGGGTCTACAGAGAGAACTGAAGAG 540
 QY 684 AGATCTGCCAAAGACATTAACCTCAGACACATCTGGAGATTTTCGGAACGCTTTCCTTC 743
 Db 541 AGATCTGCCAAAGACATTAACCTCAGACACATCTGGAGATTTTCGGAACGCTTTCCTTC 600
 QY 744 TCTTGTCTAAGGGTGACCGATCTGAGGACTTTTGGTGTGAATGAAGACTTGGCTGATTCAGA 803
 Db 601 TCTTGTCTAAGGGTGACCGATCTGAGGACTTTTGGTGTGAATGAAGACTTGGCTGATTCAGA 660
 QY 804 TCCAGGGCGCTTGTATGAGCAGAGAGAAAGGAGAGAGGAGACAGCTTAACGTTTCAA 863
 Db 661 TCCAGGGCGCTTGTATGAGCAGAGAGAAAGGAGAGAGGAGACAGCTTAACGTTTCAA 720
 QY 864 TACCATCTTACCACCAAGAGCTATCCACAACTTCGAGAGTGTTCAGAAATACACAA 923
 Db 721 TACCATCTTACCACCAAGAGCTATCCACAACTTCGAGAGTGTTCAGAAATACACAA 780

924 GTACAGTAGCATGACATGACAAAGTTCTGACCTGGAGTTGAAAGTGACATTGAGAA 983
 Db
 781 GTACAGTAGCATGACATGACAAAGTTCTGACCTGGAGTTGAAAGTGACATTGAGAA 840
 Qy
 984 ATGCGCTCACAGCTATCGTGAAGTGGCCACACAGCAAAACAGCTTTCTTTGACAGAGAGCT 1043
 Db
 841 ATGCGCTCACAGCTATCGTGAAGTGGCCACACAGCAAAACAGCTTTCTTTGACAGAGAGCT 900
 Qy
 1044 TCATCAGCCATGAAGAGTGTGGAACTGGCCATGAAGCATGATGACAGATAT-GGTTT 1102
 Db
 901 TCATC-AGCCATGAAGAGTGTGGAACTGGCCATGAAGCATGATGACAGATATGGGTTT 959
 Qy
 1103 CCGGTTCTGAATGACATGAATGATATCAAGAGCATTTCTATCAGAGATGATGATGATCT 1162
 Db
 960 CCGGTTCTGAATGACATGAATGATATCAAGAGCATTTCTATCAGAGATGATGATGATCT 1019
 Qy
 1163 -CCGTTTGCAGAGCATCTCGATGAAGCAACCAAGAGAGATGATGAGAAAATCTGTGT-GGC 1220
 Db
 1020 CCGCTTTGCCAACCTTCGGGATGAACCCAGGGAATTTTGTAAAAATCTGTGTGGC 1079
 Qy
 1221 TCTTTGTGGAGAACTAAA 1240
 Db
 1080 TCTTTGTGGAGAACTAAA 1099

RESULT 5
 AL544231
 LOCUS
 DEFINITION
 clone GS0D1019F12 5-PRIME, mRNA sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

AL544231 1183 bp mRNA linear EST 31-MAY-2003
 AL544231 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
 clone GS0D1019F12 5-PRIME, mRNA sequence.

AL544231
 AL544231.2 GI:31266076
 EST.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1183)
 Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On Feb 15, 2001 this sequence version replaced gi:12876711.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 5436.f For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0D1019DC06QPI&cluster=5436.f. Contact :
 Feng Liang Email: fliang@lifetech.com URL: Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID: CS0D1019DC06QPI.
 Location/Qualifiers
 1. .1183
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0D1019F12"
 /tissue_type="PLACENTA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /notes="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

FEATURES
 source

ORIGIN
 Query Match 66.1%; Score 1002; DB 9; Length 1183;
 Best Local Similarity 95.3%; Pred. No. 5.3e-256;
 Matches 1050; Conservative 13; Mismatches 33; Indels 6; Gaps 3;
 Qy 136 TTACAGAGAGAAATTCCTTAGTCTTTTTCAGAGAGAGTAGATGAAGACACTTTTCA 195

45 TCCGGGATGAATTTCTTTAGTTCTTTGCAAGAAGGTAGAGATAAGACACTTTTCA 104
 Qy
 196 AAAATGCAATGATATCAGAAATTCCTCAAGCAGGCTGTGTTATTGAAATGAAGAGCAG 255
 Db
 105 AAAATGCAATGATATCAGAAATTCCTCAAGCAGGCTGTGTTATTGAAATGAAGAGCAG 164
 Qy
 256 GAATATGTTCAAACTCTGAAGTCAATCAAGAGTGTCCCGGATCAGCGGTGAGCCCTAT 315
 Db
 165 GAATATGTTCAAACTCTGAAGTCAATCAAGAGTGTCCCGGATCAGCGGTGAGCCCTAT 224
 Qy
 316 CTTACCTTCAATCCATCTCGGATGTGCTGCTTCATCAAGAGCCATTAATGTTAAAGGT 375
 Db
 225 CTTACCTTCAATCCATCTCGGATGTGCTGCTTCATCAAGAGCCATTAATGTTAAAGGT 284
 Qy
 376 GTGGATGACAGACCATCTTGAATTTCTAACTTAAGCAAGAACTCAAGCGCTCAACAG 435
 Db
 285 GTGGATGACAGACCATCTTGAATTTCTAACTTAAGCAAGAACTCAAGCGCTCAACAG 344
 Qy
 436 ATCAAGCAGCATATCTCCAGGAAACAGGAAAGCCCTCGATGAAACACTTGAAGAGGCC 495
 Db
 345 ATCAAGCAGCATATCTCCAGGAAACAGGAAAGCCCTCGATGAAACACTTGAAGAGGCC 404
 Qy
 496 CTTACAGGTCACCTTGAGGAGGTTGTTTGTAGCTCTCTAACTCAAGCGCAATTTGAT 555
 Db
 405 CTTACAGGTCACCTTGAGGAGGTTGTTTGTAGCTCTCTAACTCAAGCGCAATTTGAT 464
 Qy
 556 GCTGATGAATCTCGTGTGCTGCAATGAAGGCTTTGAACTGATGAAGATCTCTAAATGAG 615
 Db
 465 GCTGATGAATCTCGTGTGCTGCAATGAAGGCTTTGAACTGATGAAGATCTCTAAATGAG 524
 Qy
 616 ATTTTGGCATCAAGACTTAACAAAGAACTCAGAGACATTAACAGGGTCTACAGAGAGGA 675
 Db
 525 ATTTTGGCATCAAGACTTAACAAAGAACTCAGAGACATTAACAGGGTCTACAGAGAGGA 584
 Qy
 676 CTGAAGAGAGATCTGCGCAAGACATAAACCCTCAGACACATCTGGAGATTTTCGGAACGCT 735
 Db
 585 CTGAAGAGAGATCTGCGCAAGACATAAACCCTCAGACACATCTGGAGATTTTCGGAACGCT 644
 Qy
 736 TTGCTTTCTCTGTAAGGTTGACCGATCTGAGGACTTTGGTGTGATGAGAGACTTTGGCT 795
 Db
 645 TTGCTTTCTCTGTAAGGTTGACCGATCTGAGGACTTTGGTGTGATGAGAGACTTTGGCT 704
 Qy
 796 GATTACAGATGCCAGGGCTTTGATGAAGCAGGAGAAAGGAGAAAGGGGACAGACGCTAAAC 855
 Db
 705 GATTACAGATGCCAGGGCTTTGATGAAGCAGGAGAAAGGAGAAAGGGGACAGACGCTAAAC 764
 Qy
 856 GTGTTCAATACCATCTCTTACCAAGAGTATCCCACTTCGAGAGTGTTCAGAAA 915
 Db
 765 GTGTTCAATACCATCTCTTACCAAGAGTATCCCACTTCGAGAGTGTTCAGAAA 824
 Qy
 916 TACACCAAGTACAGTACAGTACATGAACAAAGTTCTGGACCTGGAGTTGAAAGGTGAC 975
 Db
 825 TACACCAAGTACAGTACAGTACATGAACAAAGTTCTGGACCTGGAGTTGAAAGGTGAC 884
 Qy
 976 ATTGAGAAATGCCCTCAGCTATCGTGAAGTGGCCCAAGAGAAACAGGCTTTCTTTGCA 1035
 Db
 885 ATTGAGAAATGCCCTCAGCTATCGTGAAGTGGG-CACAGCAAAACAGGCTTTCTTTGCA 943
 Qy
 1036 GAGAGCTTCTCAAGCCATGAAGGTTGTGAACTCGCCATGAAGCATTGATCAGGATT 1095
 Db
 944 GAGAGCTTCTCAAGCCATGAAGGTTGTGAACTCGCCATGAAGCATTGATCAGGATT 1003
 Qy
 1096 ATGTTTCCCGTCTCGAAATTTGACATGAATGATCAAAAGCATTTCTATCAGAGAGTAT 1155
 Db
 1004 ATGTTTCCCGTCTCGAAATTTGACATGAATGATCAAAAGCATTTCTATCAGAGAGG--GTA 1061
 Qy
 1156 GGTATCTCCCTTTCGCAAGCCATCTCGATGAACCAAGAGAGATGATGAGAAATCTCTG 1215
 Db
 1062 TGGWWYTCCTTTGCGCAAGCMWYCTGG---ATGAACCAAGAGGATATGAGAAATCTGTG 1118
 Qy
 1216 GTGGCTCTCTTGTGGAGGAAACT 1237

Db 1119 GGCTCTTKTGGGGAATWAAAT 1140

RESULT 6
BU902298 1163 bp mRNA linear EST 17-OCT-2002
LOCUS AGENCOURT_10127317 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5503136
DEFINITION 5', mRNA sequence.
ACCESSION BU902298
VERSION BU902298.1 GI:24084211
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1163)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM14060 Row: f Column: 01
High quality sequence stop: 821.
Location/Qualifiers
1..1163
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6503136"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 71"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dh. Average insert size 2.1 kb. "
FEATURES
source

Query Match 65.2%; Score 988.6; DB 13; Length 1163;
Best Local Similarity 95.2%; Pred. No. 2e-252;
Matches 1063; Conservative 0; Mismatches 49; Indels 5; Gaps 4;
135 CTTGAGAGAGAAATTCCTTTAGTCTTTGCAAGAGAGGTAGAGATAAAGACA-CCTTTT 193
24 CTTGAGAGAGAAATTCCTTTAGTCTTTGCAAGAGAGGTAGAGATAAAGACANCTTTT 83
194 CAAAAATGGCAATGGTATCAGAAATTCCTCAAGAGGCGCTGTTTATGAAATGAGAGC 253
84 CAAAAATGGCAATGGTATCAGAAATTCCTCAAGAGGCGCTGTTTATGAAATGAGAGC 143
254 AGGAATATGTTCAAACTGTGAAGTCATCCAAAGGTGTCGCGGATCAGCGGTGAGCCCT 313
144 AGGAATATGTTCAAACTGTGAAGTCATCCAAAGGTGTCGCGGATCAGCGGTGAGCCCT 203
314 ATCTTACCTTCAATCCATCCCTCGGATGTCGCTTGCATTAAGGCCATATGGTTAAAG 373
204 ATCTTACCTTCAATCCATCCCTCGGATGTCGCTTGCATTAAGGCCATATGGTTAAAG 263
374 GTGTGGATGAGCAACCATATTCACATTCCTAACTAAGCGAAACAAATGCACAGCTCAAC 433
264 GTGTGGATGAGCAACCATATTCACATTCCTAACTAAGCGAAACAAATGCACAGCTCAAC 323
434 AGATCAAAAGCAGCATATCTCCAGAAACAGAAAGCCCTGGATGAAACACTGAGAAAG 493
324 AGATCAAAAGCAGCATATCTCCAGAAACAGAAAGCCCTGGATGAAACACTGAGAAAG 383
494 CCCCTACAGGTACCTTGGAGAGGTTGTTTAGCTCTGCTTAAAACTCCAGCGCAATTTG 553

Db 384 CCCTTACAGGTACCTTGGAGAGGTGTTTGTAGCTCTGCTTAAAACTCCAGCGCAATTTG 443
QY 554 ATGCTGATGAATCTTGGTCTGCCATGAAGGCGCTTGAACCTGATGAAGATACTCTAAATTTG 613
Db 444 ATGCTGATGAATCTTGGTCTGCCATGAAGGCGCTTGAACCTGATGAAGATACTCTAAATTTG 503
QY 614 AGATTTTGGCATCAAGACTAAACAAAGAAATCAGAGACATTAACAGGGTCTTACAGAGAGG 673
Db 504 AGATTTTGGCATCAAGACTAAACAAAGAAATCAGAGACATTAACAGGGTCTTACAGAGAGG 563
QY 674 AACTGAGAGAGATCTGGCCAAAGACATACCTCAGACACATCTGGAGATTTTCGGAACG 733
Db 564 AACTGAGAGAGATCTGGCCAAAGACATACCTCAGACACATCTGGAGATTTTCGGAACG 623
QY 734 CTTTGTCTTCTTCTTGAAGGTCAGCGATCTGAGGCTTTGGTGTGAATGAAGACTTTGG 793
Db 624 CTTTGTCTTCTTCTTGAAGGTCAGCGATCTGAGGCTTTGGTGTGAATGAAGACTTTGG 683
QY 794 CTGATTCAGATGCCAGGCGCTTGTATGAAGCAGGAGAAAGAGAAAGGGGACAGACGTAA 853
Db 684 CTGATTCAGATGCCAGGCGCTTGTATGAAGCAGGAGAAAGAGAAAGGGGACAGACGTAA 743
QY 854 AGTGTTCATACCATCTTACACAGAGAGCTATCCACACTTCGCGAGAGTGTTCAGA 913
Db 744 AGTGTTCATACCATCTTACACAGAGAGCTATCCACACTTCGCGAGAGTGTTCAGA 803
QY 914 AATACACCAAGTACAGTAAGCATGACATGAACAAAGTTCTGGACCTGGAGTTGAAAGTG 973
Db 804 AATACACCAAGTACAGTAAGCATGACATGAACAAAGTTCTGGACCTGGAGTTGAAAGTG 863
QY 974 ACATTTGAGAAATGCTTACAGCTATCTGAGTGGCGGACCAAGCAAAACCGACTTTCTTTG 1033
Db 864 ACATTTGAGAAATGCTTACAGCTATCTGAGTGGCGGACCAAGCAAAACCGACTTTCTTTG 923
QY 1034 CAGAGAAGCTTCATCAAGCCATGAAGGTGTGGAACCTCGCCATAAGGCAATGATCAGGA 1093
Db 924 CAGAGAAGCTTCATCAAGCCATGAAGGTGTGGAACCTCGCCATAAGGCAATGATCAGGG 983
QY 1094 TTATGGTTTCCGTTCTGAAATGACATGAATGATAT-CAAAGCATTTCTATCAGAG-AT 1151
Db 984 ATATGGTTCCTCGTTCTGAAATGACATTTGAATGAATCAAGGCTTTCTATCAAAAGAG 1043
QY 1152 GTATGGTATCTCCCTTTGCCAAGCCATCTCTGGAT--GAAACCAAGGAGAGATGATGAGAAA 1209
Db 1044 GAATGGAATCTCCCTTTGGCACACCCCTCTGTTGGAACCAAGGGAATATGGAATA 1103
QY 1210 ATCTTGGTGGCTTCTTTGGAGGAAACTTAAACATTC 1246
Db 1104 AATCCCGGGGCGCTCTTTGTGGGAAGAAAAAATAC 1140

RESULT 7
AL553114 1201 bp mRNA linear EST 31-MAY-2003
LOCUS AL553114 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS01074YC03 5-PRIME, mRNA sequence.
ACCESSION AL553114
VERSION AL553114.2 GI:31274928
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1201)
AUTHORS Li W.B., Gruber C., Jessee J. and Polayes D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 15, 2001 this sequence version replaced gi:12892649.
Contact: Genoscope
Genoscope - Centre National de Sequenage
Bp 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 5436.f For more information about this cluster, see <http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0D1074AB020P1&cluster=5436.f>. Contact : Peng Liang Email : fliang@lifetech.com URL : Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0D1074AB020P1.

FEATURES

Location/Qualifiers
1. .1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1074AB020P1"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 65.1%; Score 986.8; DB 9; Length 1201;
Best Local Similarity 99.3%; Pred. No. 6e-252;
Matches 1011; Conservative 1; Mismatches 3; Indels 3; Gaps 2;

QY 17 AGCTGAGCCATGTTCTTCAACAGAGGCGACCAATTAACCTTCTGGTGTCTAGGTGT 76
DB 63 ATCTGAGCCATGTTCTTCAACAGAGGCGACCAATTAACCTTCTGGTGTCTAGGTGT 122

QY 77 GGCTTCTTTAAATCTTATTAATCAAGAGCCCAAGTCTCCAGTCCAGGTGAAATCT 136
DB 123 GGCTTCTTTAAATCTTATTAATCAAGAGCCCAAGTCTCCAGTCCAGGTGAAATCT 182

QY 137 TCAGAGAGATTTCTTTAGTCTTTGCAAGAGGTAGATGATGATGATGATGATGATGAT 196
DB 183 TCAGAGAGATTTCTTTAGTCTTTGCAAGAGGTAGATGATGATGATGATGATGATGAT 242

QY 197 AAATGGCAATGGTATCAAGATTCCTCAAGAGGCGTGGTGTATTAAGAAAGAGGAGG 256
DB 243 AAATGGCAATGGTATCAAGATTCCTCAAGAGGCGTGGTGTATTAAGAAAGAGGAGG 302

QY 257 ATATGTTCAACTGTGAAGTCAATCAAGAGTGTGTCGGATGAGCGTCAAGGCTATC 316
DB 303 ATATGTTCAACTGTGAAGTCAATCAAGAGTGTGTCGGATGAGCGTCAAGGCTATC 362

QY 317 CTACCTTCAATCCATCTCGATGTCGCTGCTTGCATAGAGCCCATATGTTAAAGGTG 376
DB 363 CTACCTTCAATCCATCTCGATGTCGCTGCTTGCATAGAGCCCATATGTTAAAGGTG 422

QY 377 TGGATGAAGCAACCATATGATGATTTCAATCAAGCAAAATGCAAGGCTCAACAGA 436
DB 423 TGGATGAAGCAACCATATGATGATTTCAATCAAGCAAAATGCAAGGCTCAACAGA 482

QY 437 TCAGAGAGCATATCTCCAGAGAAACAGAGAGCCCTCGATGAAACACTGAAGAAAGCCC 496
DB 483 TCAGAGAGCATATCTCCAGAGAAACAGAGAGCCCTCGATGAAACACTGAAGAAAGCCC 542

QY 497 TTACAGGTCACTTGGAGGAGTGTGTTAGTCTCTCTAAAACCTCCAGGCAATTTGATG 556
DB 543 TTACAGGTCACTTGGAGGAGTGTGTTAGTCTCTCTAAAACCTCCAGGCAATTTGATG 602

QY 557 CTGATGAATCTGCTGCTGATGAGGCGCTTGGAACTGATGATGATGATGATGATGATGAT 616
DB 603 CTGATGAATCTGCTGCTGATGAGGCGCTTGGAACTGATGATGATGATGATGATGATGAT 662

QY 617 TTTTGGCATCAAGAACTAACAAAGAAATCAGAGACATTAACAGGCTCTACAGAGGAGAC 676
DB 663 TTTTGGCATCAAGAACTAACAAAGAAATCAGAGACATTAACAGGCTCTACAGAGGAGAC 722

QY 677 TGAAGAGAGATCTGGCAAGAGCATTAACCTCAGACACATCTGGAGATTTTCGGAGCGTT 736
DB 723 TGAAGAGAGATCTGGCAAGAGCATTAACCTCAGACACATCTGGAGATTTTCGGAGCGTT 782

QY 737 TCCTTTCTTTGCTTAAGGGTGACCGATCTGAGGACTTTTGGTGTGATGAGAGCTTGGCTG 796
DB 783 TCCTTTCTTTGCTTAAGGGTGACCGATCTGAGGACTTTTGGTGTGATGAGAGCTTGGCTG 842

QY 797 ATTCAAGTGCAGAGGCGCTTGTATGAAGCAGAGAGAAAGGAGAAAGGGGACAGACGTAACG 856
DB 843 ATTCAAGTGCAGAGGCGCTTGTATGAAGCAGAGAGAAAGGAGAAAGGGGACAGACGTAACG 902

QY 857 TGTTCATTAACCATCTTACCACAGAGCTATCCAACTTCGCGAGAGTGTTCAGAAAT 916
DB 903 TGTTCATTAACCATCTTACCACAGAGCTATCCAACTTCGCGAGAGTGTTCAGAAAT 962

QY 917 ACACCAAGTACAGTAAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 976
DB 963 ACACCAAGTACAGTAAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1022

QY 977 TTGAGAAATGCTTCACAGCTATCGTGAAGTGGCCACAGCAAAACAGCTTCTTTTCG 1034
DB 1023 TTGAGAAATG-CTCACAGCTATCGTGAAGTGGCCACAC--GCAACCAAGCTTCTTTTCG 1077

RESULT 8

AL575425/c
LOCUS AL575425 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0D1060YMI3 3-PRIME, mRNA sequence.
ACCESSION AL575425
VERSION AL575425
KEYWORDS EST.
SOURCE AL575425.2 GI:31313733
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1201)
AUTHORS Li W.B., Gruber C., Jessee J. and Polayes D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 16, 2001 this sequence version replaced gi:12936576.
Contact: Genoscope
Genoscope - Centre National de Sequencage
Bp 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 5436.f For more information about this cluster, see <http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0D1060AG07NP1&cluster=5436.f>. Contact : Peng Liang Email : fliang@lifetech.com URL : <http://fulllength.invitrogen.com/> Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0D1060AG07NP1.

FEATURES

Location/Qualifiers
1. .1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1060YMI3"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 65.1%; Score 986.8; DB 9; Length 1201;
Best Local Similarity 96.3%; Pred. No. 6e-252;
Matches 989; Conservative 11; Mismatches 21; Indels 0; Gaps 0;

QY 488 AGAAGCCCTTACAGGTCCAGCTTGGAGGAGTGTGTTTAGCTCTGCTTAAACACTCCAGCG 547
DB 1022 AAAAGCCCTTACAGGTCCAGCTTGGAGGAGTGTGTTTAGCTCTGCTTAAACACTCCAGCG 963

	QY	548	AATTTGATCGTGATGAACCTTCGTGCTGCCATGAAGGGCCCTTGGAACTGTGATGAGAATACTC	607
	Db	962	AAATTTGATCGTGATGAACCTCCGGCTSCCAATAAAGGGCCCTTGGAACTGTGATGAGAATACTC	903
	QY	608	TAAATTGAGATTTTTGGCATCAAGAACAATAACAAAGAAAATCACAGACATTAACAGGGTCTACA	667
	Db	902	TAAATTGAGATTTTTGGCATCAAGAACAATAACAAAGAAAATCACAGACATTAACAGGGTCTACA	843
	QY	668	GAGAGGAACTGAGAGAGATCTGGCCAAAGACATACCTCAGACACATCTGGAGATTTTC	727
	Db	842	GAGAGGRACCTGAAGAGAGATCTGGCCAAAGACATACCTCAGACACATCTGGAGATTTTC	783
	QY	728	GGAACCGCTTTGCCTTTCTCTTGGTCAAGGGTACCAGATCTGAGGACTTTGGTGTGAATCAAG	787
	Db	782	GGAACCGCTTTGCCTTTCTCTTGGTCAAGGGTACCAGATCTGAGGACTTTGGTGTGAATCAAG	723
	QY	788	ACTTGGCTGANTCAGATGCCAGGGCCTTGATGAAAGCAGGAGAAAGAGGAAAGGGGACAG	847
	Db	722	ACTTGGCTSATTCAGATGCCAGGGCCTTGATGAAAGCAGGAGAAAGAGGAAAGGGTACAG	663
	QY	848	ACGTAAACGTGTTCAATACCATCTTACCACCAAGAGCTATCCACAACCTTCGCAGAGTGT	907
	Db	662	ACGTAAACGTGTTCAATACCATCTTACCACCAAGAGCTATCCACAACCTTCGCAGAGTGT	603
	QY	908	TTCAGAAAAATACACCAAGTACAGTAAGCATGACATGAACAAAGTTCTGGACCTCGAGTTGA	967
	Db	602	TTCAGAAAAATACACCAAGTACAGTAAGCATGACATGAACAAAGTTCTGGACCTCGAGTTGA	543
	QY	968	AAGGTGACATTTGAGAAATGCTCCACAGCTATCGTGAAGTCGGCCAACGAACAAACCAAGCTT	1027
	Db	542	AAGGTGACATTTGAGAAATGCTCCACAGCTATCGTGAAGTCGGCCAACGAACAAACCAAGCTT	483
	QY	1028	TCCTTTGCAGAGAAAGCTTCATCAAGCCATGAAGAGTGTGGAATCTCGGCCATAAAGGCATTGA	1087
	Db	482	TCCTTTGCAGAGAAAGCTTCATCAAGCCATGAAGAGTGTGGAATCTCGGCCATAAAGGCATTGA	423
	QY	1088	TCAGGATTAATGGTTTCCCGTTCTGAAATTGACATGAATGATATCAAAGCATTTCTATCAGA	1147
	Db	422	TCAGGATTAATGGTTTCCCGTTCTGAAATTGACATGAATGATATCAAAGCATTTCTATCAGA	363
	QY	1148	AGATGTATGGTATCTCCCTTTGCCAAGCCATCTCTGGATGAAACCAAGGAGAGATGATGAGA	1207
	Db	362	AGATGTATGGTATCTCCCTTTGCCAAGCCATCTCTGGATGAAACCAAGGAGAGATGATGAGA	303
	QY	1208	AAATCTCTGGTGGCTTTTGTGGAGGAAACATAACATTTCCCTTGATGTCTCAAGCTATGA	1267
	Db	302	AAATCTCTGGTGGCTTTTGTGGAGGAAACATAACATTTCCCTTGATGTCTCAAGCTATGA	243
	QY	1268	TCAGAGACATTTAAATATATATTTTCATCTAAGCTTAAATAGGAAAGTTTCTTCAAC	1327
	Db	242	TCAGAGACATTTAAATATATATTTTCATCTAAGCTTAAATAGGAAAGTTTCTTCAAC	183
	QY	1328	AGGATTCAGTGTAGCTACCTACATGCTGAAATAATATAGCCTTTAAATCAATTTTATATT	1387
	Db	182	AGGATTCAGTGTAGCTACCNACATGCTGAAAAATATAGCCTTTAAATCAATTTTATATT	123
	QY	1388	ATAACTCTGATTAATAGAGATAAGTCCATTTTAAAAAATGTTTCCCCCAACCATAAAA	1447
	Db	122	ATAACTCTGATTAATAGAGATAAGTCCATTTTAAAAATGTTTCCCCCAACCATAAAA	63
	QY	1448	CCCTATACAAGTTGTTCTAGTAAACAATACATGAGAAAGATGCTCTATGTAGCTGAAATAAA	1507
	Db	62	CCCTATACAAGTTGTTCTAGTAAACAANACATGAGAAAGAGTCTTAGKRNNGAARNAAA	3
	QY	1508	A 1508	
	Db		2 A 2	

RESULT 9
AL541874
LOCUS
DEFINITION AL541874 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE007YN11
EST 12-MAY-2003 linear

5-PRIME, mRNA sequence.
AL541874
AL541874.2 GI:30546467
EST.
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
Li W.B., Gruber C., Jessee J. and Polayes D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 15, 2001 this sequence version replaced gi:12873366.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 5436.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DE007CG06QPl&cluster=5436.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DE007CG06QPl.
FEATURES
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE007YN11"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
ORIGIN
Query Match 64.9%; Score 984; DB 9; Length 1201;
Best Local Similarity 99.0%; Pred. No. 3.4e-251;
Matches 990; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 150 TCCTCTTAGTCTCTTCGACGAGGTAGAGTAAAGACACTTTTCAAATATGCAATGGT 209
Db 62 TCCTCTTAGTCTCTTCGACGAGGTAGAGTAAAGACACTTTTCAAATATGCAATGGT 121
QY 210 ATCAGAATTCTCTCAAGCAGGCGCTGGTTATTATGAAATGAAGACGAGGAATATGTTCAAC 269
Db 122 ATCAGAATTCTCTCAAGCAGGCGCTGGTTATTATGAAATGAAGACGAGGAATATGTTCAAC 181
QY 270 TGTGAAGTCATCAAAAGGTGTTCCCGGATCAGCGGTGAGCCCTTCTCTACCTTCAATCC 329
Db 182 TGTGAAGTCATCAAAAGGTGGTCCCGATCAGCGGTGAGCCCTTCTCTACCTTCAATCC 241
QY 330 ATCCTCGGATGTCGCTGCCTTGCATAAGGCCATATGTTTAAAGGTGTGATGAAGCAAC 389
Db 242 ATCCTCGGATGTCGCTGCCTTGCATAAGGCCATATGTTTAAAGGTGTGATGAAGCAAC 301
QY 390 CATCATGTGATCTTAACATAAGCGAACCAATGCACAGCGTCAACAGATCAAGACGAGATA 449
Db 302 CATCATGTGATCTTAACATAAGCGAACCAATGCACAGCGTCAACAGATCAAGACGAGATA 361
QY 450 TCTCCAGGAAACAGGAAAGCCCTTGGATGAAACACTGAAGAAAGCCCTTACAGTCACT 509
Db 362 TCTCCAGGAAACAGGAAAGCCCTTGGATGAAACACTGAAGAAAGCCCTTACAGTCACT 421
QY 510 TGAGGAGTTGTTTATGCTCTGCTTAAACCTCCAGCGCAATTTGATCGCTGATGAACCTCG 569
Db 422 TGAGGAGTTGTTTATGCTCTGCTTAAACCTCCAGCGCAATTTGATCGCTGATGAACCTCG 481
QY 570 TGCTGCCATGAAGGGCCTTGGAACTGATGAAGATACCTTAATTGAGATTTTGGCATCAAG 629

RESULT 9
AL541874
LOCUS
DEFINITION

```
Db 482 TGCTGCCATGAAGGCCCTTGGAACTGATGAAGATACCTAATTGAGATTTTGGCATCAAG 541
Qy 630 AACTAAACAAGAAATCAGACACATTAAACAGGGTCTACAGAGAGAACTGAAGAGAGATCT 689
Db 542 AACTAAACAAGAAATCAGACACATTAAACAGGGTCTACAGAGAGAACTGAAGAGAGATCT 601
Qy 690 GGCCAAAGACATAAAGCTCAGACACATCTGGAGATTTTCGGAACGCTTTCCTCTTGC 749
Db 602 GGCCAAAGACATAAAGCTCAGACACATCTGGAGATTTTCGGAACGCTTTCCTCTTGC 661
Qy 750 TAAGGGTCAACGATCTGAGGACTTTGGTGTGAATGAAGACTTGGCTGATTCAGATGCCAG 809
Db 662 TAAGGGTCAACGATCTGAGGACTTTGGTGTGAATGAAGACTTGGCTGATTCAGATGCCAG 721
Qy 810 GGCCTTGTATGAAGCAGAGAGAAAGAGAAAGGGGACAGACGTAACAGTGTTCATATACAT 869
Db 722 GGCCTTGTATGAAGCAGAGAGAAAGAGAAAGGGGACAGACGTAACAGTGTTCATATACAT 781
Qy 870 COTTAACCAAGAGCTATCCACAACTTTCGAGAGTGTTCAGAAATACACCAAGTACAG 929
Db 782 COTTAACCAAGAGCTATCCACAACTTTCGAGAGTGTTCAGAAATACACCAAGTACAG 841
Qy 930 TAAGCATGACATGAACAAAGTTCTGCACTGAGCTGAAAGTGAAGTGAAGTGAAGTGAAGT 989
Db 842 TAAGCATGACATGAACAAAGTTCTGCACTGAGCTGAAAGTGAAGTGAAGTGAAGTGAAGT 901
Qy 990 CACAGCTATCGTGAAGTGGCCACAGCAACACCAAGCTTCTTCGAGAGAGCTTCATCA 1049
Db 902 CACAGCTATCGTGAAGTGGCCACAGCAACACCAAGCTTCTTCGAGAGAGCTTCATCA 961
Qy 1050 AGCCATGAAGTGTGGAACCTCGCCATGAAGCAATGATCAGGATTAATGTTTCCCGTTC 1109
Db 962 AGCCATGAAGTGTGGAACCTCGCCATGAAGCAATGATCAGGATTAATGTTTCCCGTTC 1021
Qy 1110 TGAATTCATGATGATATCAAGCAATTCATCAGAG 1149
Db 1022 TGAATTCATGATGATATCAAGCAATTCATCAGAGATG 1061
```

```
RESULT 10
LOCUS AL574991/c 1136 bp mRNA linear EST 01-JUN-2003
DEFINITION AL574991 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1064YE07 3-PRIME, mRNA sequence.
ACCESSION AL574991
VERSION AL574991.2 GI:31313300
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 16, 2001 this sequence version replaced gi:12935721.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 5436.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0D1064AC04NP1&cluster=5436.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0D1064AC04NP1.
FEATURES
source
1..1136
/organism="Homo sapiens"
/mol_type="mRNA"
```

```
/db_xref="taxon:9606"
/clone="CS0D1064YE07"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 64.1%; Score 971.4; DB 9; Length 1136;
Best Local Similarity 93.8%; Pred. No. 7.7e-248;
Matches 1027; Conservative 15; Mismatches 46; Indels 7; Gaps 3;

Qy 396 TGACATTTCTAACTAAGCGAAACAATGCACAGCGTCAACAGATCAAGACAGCATATCTCCA 455
Db 1088 TGAGCACTTATACATTTAAAAAAXAXAXAXAXAXAXAXAXAXAXAXAXAXAXAXAXAX 1029
Qy 456 GGAACAGAGAAAGCCCTGGATGAACACTGAAGAAAGCCCTTACAGGTACCTTGAGGA 515
Db 1028 MGAAACAGAGAAAGCCCTGATRAACACAKAAGAAAR-----SCCTACCGGTCACTTGA 974
Qy 516 GCTTGTCTTGTAGCTCTGCTAAAGTCTCCAGCGCAATTTGATGCTGATGAACCTTCGTC 575
Db 973 -GTGTGTTTGTAGCTCTGCTAAAGTCTCCAGCGCAATTTGATGCTGATGAAC-TCGTGCT 916
Qy 576 CATGAAGGGGCTTTGGAACTGATGAAGATACTCTAATGAGATTTTGGCATCAAGAACTAA 635
Db 915 CATGAAGGGGCTTTGGAACTGATGAAGATACTCTAATGAGATTTTGGCATCAAGAACTAA 856
Qy 636 CAAGAATTCAGAGACATTAACAGGGTCTACAGAGAGGAACCTGAAGAGAGATCTGCCCAA 695
Db 855 CAAGAATTCAGAGACATTAACAGGGTCTACAGAGAGGAACCTGAAGAGAGATCTGCCCAA 796
Qy 696 AGACATAAAGCTCAGACACATCTGGAGATTTTCGGAACGCTTTCTCTTCTTCTTCTTCT 755
Db 795 AGACATAAAGCTCAGACACATCTGGAGATTTTCGGAACGCTTTCTCTTCTTCTTCTTCT 736
Qy 756 TGACCATCTGAGGACTTTGGTGTGAATGAAGACTTTGGCTGATTCAGATGCCAGGSCCTT 815
Db 735 TGACCATCTGAGGACTTTGGTGTGAATGAAGACTTTGGCTGATTCAGATGCCAGGSCCTT 676
Qy 816 GTATGAAGCAGGAGAAAGGAGAAAGGGGACAGACGTTAAACGTTTCAATACCATCTTAC 875
Db 675 GTATGAAGCAGGAGAAAGGAGAAAGGGGACAGACGTTAAACGTTTCAATACCATCTTAC 616
Qy 876 CACCAAGCTATCCCAACTTCGAGAGTGTTCGAGAAATACACCAAGTACAGTAAAGCA 935
Db 615 CACCAAGCTATCCCAACTTCGAGAGTGTTCGAGAAATACACCAAGTACAGTAAAGCA 556
Qy 936 TGACATGAACAAAGTTCTGGACCTGAGTGAAGGTGAAGGTGACATTGAGAAATGCCCTCAG 995
Db 555 TGACATGAACAAAGTTCTGGACCTGAGTGAAGGTGACATTGAGAAATGCCCTCAG 496
Qy 996 TATCGTGAAGTGGCCACAAAGCAAAACAGCTTTCTTTTTCGAGAGAGCTTCATCAAGCCAT 1055
Db 495 TATCGTGAAGTGGCCACAAAGCAAAACAGCTTTCTTTTTCGAGAGAGCTTCATCAAGCCAT 436
Qy 1056 GAAAGTGTGGAACCTGCCATGAAGCATTTGATGAGATTTATGGTTTCCCGTTCTGAAAT 1115
Db 435 GAAAGTGTGGAACCTGCCATGAAGCATTTGATGAGATTTATGGTTTCCCGTTCTGAAAT 376
Qy 1116 TGACATGAATGATATCAAGCAATTTCTATCAGAGATGATATGATATCTCCCTTTCGCAAG 1175
Db 375 TGACATGAATGATATCAAGCAATTTCTATCAGAGATGATATGATATCTCCCTTTCGCAAG 316
Qy 1176 CATCCTGGATGAACCAAGAGAGATGATGAGAAATCTCGTGGCTCTTTTGTGGAGGAAA 1235
Db 315 CWTCTGGATGAACCAAGAGAGATTTATGAGAAATCTCGTGGCTCTTTTGTGGAGGAAA 256
Qy 1236 CTAAACATCTCCCTTGTATGGTCTCAAGCTATGATCAGAGACTTTAATATATATTTTCAT 1295
Db 255 CTAAACATCTCCCTTGTATGGTCTCAAGCTATGATCAGAGACTTTAATATATATTTTCAT 196
```

```

QY 1296 CCTATAAGCTTAAATAGGAAGTTTCTTCAACAGGATTACAGTGTAGCTACCTACATGCT 1355
Db 195 CCTATAAGCTTAAATAGGAAGTTTCTTCAACAGGATTACAGTGTAGCTACCTACATGCT 136
QY 1356 GAAATAATATAGCTTTAAATCTTTTATATATATACTCTGTATATAGAGATAAGTCCA 1415
Db 135 GAAATAATATAGCTTTAAATCTTTTATATATATACTCTGTATATAGAGATAAGTCCA 76
QY 1416 TTTTAAAAATGTTTTCCTCCCAACCAATAAACCTATACAGTTGTTCTAGTAAACAATA 1475
Db 75 TTTTAAAAATGTTTTCCTCCCAACCAATAAACCTATACAGTTGTTCTAGTAAACAATA 16
QY 1476 CATGAGAAAGATGTC 1490
Db 15 CATGAGAAAGATGTC 1

RESULT 11
AL553129 1201 bp mRNA linear EST 31-MAY-2003
LOCUS
DEFINITION
AL553129 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
Clone CSOD10741116 5-PRIME, mRNA sequence.
ACCESSION
AL553129
VERSION
AL553129.2 GI:31274943
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
Li, W.-B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 15, 2001 this sequence version replaced gi:12892678.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 5436.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CSOD1074BE080P1&cluster=5436.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CSOD1074BE080P1.
FEATURES
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSOD1074Y116"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-ciligo (dt)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 63.3%; Score 959.8; DB 9; Length 1201;
Best Local Similarity 95.5%; Pred. No. 9.7e-245;
Matches 101; Conservative 8; Mismatches 35; Indels 5; Gaps 3;

QY 142 GAAGAATTTCTTTAGTCTTTTGAAGAAGGTAGAGATAAGACATTTTCAAAAATG 201
Db 66 GAAGAATTTCTTTAGTCTTTTGAAGAAGGTAGAGATAAGACATTTTCAAAAATG 125
QY 202 GCAATGGTATCAGATTCTCTCAAGCAGCGCTGTTTATTGAAATGAAGACAGGATAT 261
Db 126 GCAATGGTATCAGATTCTCTCAAGCAGCGCTGTTTATTGAAATGAAGACAGGATAT 185

```

```

QY 262 GTTCAAACTGTGAAGTTCATCCAAAGTGG-TCCGGATCAGCGGTGAGCCCTATCTCTAC 320
Db 186 GTTCAAACTGTGAAGTTCATCCAAAGTGGATCCGGATCAGCGGTGAGCCCTATCTCTAC 245
QY 321 CTTCAATCCATCTCTCGGATGTCTGCTGCTTGCATATAAGGCCATTAATGGTTAAAGGTGTGA 380
Db 246 CTTCAATCCATCTCTCGGATGTCTGCTGCTTGCATATAAGGCCATTAATGGTTAAAGGTGTGA 305
QY 381 TGAAGCAACCATCATTTGACATTTCTAACTAAGCGAAACAATGCACAGCGTCAACAGATCAA 440
Db 306 TGAAGCAACCATCATTTGACATTTCTAACTAAGCGAAACAATGCACAGCGTCAACAGATCAA 365
QY 441 AGCAGCATATCTCCAGAAAACAGAAAAGCCCTTGGATGAAACACTGAAAGAACCCCTTAC 500
Db 366 AGCAGCATATCTCCAGAAAACAGAAAAGCCCTTGGATGAAACACTGAAAGAACCCCTTAC 425
QY 501 AGGTCACTTTGAGGAGGTTGTTTTAGCTCTGTCTTAAAAAACTCCAGCGCAATTTGATGCTGA 560
Db 426 AGGTCACTTTGAGGAGGTTGTTTTAGCTCTGTCTTAAAAAACTCCAGCGCAATTTGATGCTGA 485
QY 561 TGAATTCGTGCTGCCATGAGGGCCCTTGGAACTGATGAAGTACTCTAATTTGAGATTTT 620
Db 486 TGAATTCGTGCTGCCATGAGGGCCCTTGGAACTGATGAAGTACTCTAATTTGAGATTTT 545
QY 621 GGCATCAAGAACTAAACAAGAAATTCAGACACATTAACAGGGTCTACAGAGAGAACTGAA 680
Db 546 GGCATCAAGAACTAAACAAGAAATTCAGACACATTAACAGGGTCTACAGAGAGAACTGAA 605
QY 681 GAGAGATCTGGCCAAAGACATTAACCTCAGACACATCTGGAGATTTTCGGAACGCTTGTCT 740
Db 606 GAGAGATCTGGCCAAAGACATTAACCTCAGACACATCTGGAGATTTTCGGAACGCTTGTCT 665
QY 741 TTCTCTTTGTAAGGGTGACCGATCTGAGGACTTTGGTGTGAATGAAGACTTGGCTGATTTC 800
Db 666 TTCTCTTTGTAAGGGTGACCGATCTGAGGACTTTGGTGTGAATGAAGACTTGGCTGATTTC 725
QY 801 AGATGCCAGGGCTTGTATGAGCAGGAGAAAGGAGAAAGGGACAGAGCTAAACGTGT 860
Db 726 AGATGCCAGGGCTTGTATGAGCAGGAGAAAGGAGAAAGGGACAGAGCTAAACGTGT 785
QY 861 CAATACCACCTTTACCAACAGAGACTATCCAACTTCGACAGAGTGTTCAGAAATATACAC 920
Db 786 CAATACCACCTTTACCAACAGAGACTATCCAACTTCGACAGAGTGTTCAGAAATATACAC 845
QY 921 CAATCAGTAAGCATGACATGAACAAGTTCTGGACCTGGAGTTCGAAAGGTGACATTCGA 980
Db 846 CAATCAGTAAGCATGACATGAACAAGTTCTGGACCTGGAGTTCGAAAGGTGACATTCGA 905
QY 981 GAAATGCTCAGCTATCGTGAAGTGGCCCAACAGCAAAACCAAGCTTTCTTTGCAGAGAA 1040
Db 906 GAAATGCTCAGCTATCGTGAAGTGGCCCAACAGCAAAACCAAGCTTTCTTTGCAGAGAA 965
QY 1041 GCTTCATCAAGCCATGAAGGTGTGGAACTGCCATAGGCAATTCATCAGGATTTATGTT 1100
Db 966 GCTTCATCAAGCCATGAAGGTGTGGAACTGCCATAGGCAATTCATCAGGATTTATGTT 1025
QY 1101 TTCGGTCTGAAATTCAGATGATGATCAAGCATTTCTATCAGAAAGATGTATGGTAT 1160
Db 1026 TTCGGTCTGAAATTCAGATGATGATCAAGCATTTCTATCAGAAAGATGTATGGTAT 1081
QY 1161 CTCCTTTGCCAGCCATCTCTGGATGAAACCAAGGAGA 1199
Db 1082 CTCCTTTGCCAGCCATCTCTGGATGAAACCAAGGATGAGA 1120

RESULT 12
AL544959 1201 bp mRNA linear EST 31-MAY-2003
LOCUS
DEFINITION
AL544959 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
Clone CSOD1012P05 5-PRIME, mRNA sequence.
ACCESSION
AL544959
VERSION
AL544959.2 GI:31266800
KEYWORDS
EST.

```

544 GAGATTTTGGCATCAAGAAGCTAACAAAGAAATCATGAGACATTAAACGGGTTCTACAGAGAG 603

673 GAACTGAAGAGAGATCTGGCCAAAGACATAACTCTCAGACACATCTGAGAGATTTTCGGAAC 732

604 GAACTGAAGAGAGATCTGGCCAAAGACATAACTCTCAGACACATCTGAGAGATTTTCGGAAC 663

733 GCTTTGCTTTCTCTGCTAAGGGTGACCGATCTGAGGAGCTTTGGTGTGAATGAAGAAGACTTG 792

664 GCITTTGCTTTCTCTGCTAAGGGTGACCGATCTGAGGAGCTTTGGTGTGAATGAAGAAGACTTG 723

793 GCTGATTTCAGATGCCAGGGGCTTTGTTATGAAGCAGAGAGAAAGGAGAAAGGGGACAGAGCTA 852

724 GCTGATTTCAGATGCCAGGGGCTTTGTTATGAAGCAGAGAGAAAGGAGAAAGGGGACAGAGCTA 783

853 AACGTGTTTCAATACCATCTTTACCACCAAGAGCTATCCACAACCTTCGCAGAGAGTGTTCAG 912

784 AACGTGTTTCAATACCATCTTTACCACCAAGAGCTATCCACAACCTTCGCAGAGAGTGTTCAG 843

913 AAATACACCAAGTACAGTAAAGCATGACATGACAAAGTCTTGGACCTGGAG-TTGAAGG 971

844 AAATACACCAAGTACAGTAAAGCATGACATGACAAAGTCTTGGACCTGGAGTTTGAAGG 903

972 TGACATTTGAGAAATGCGCTCACAGCTATCGTGAAGTCGCCCAACAGCAACACGCTTTCTT 1031

904 TGACATTTGAGAAATGCGCTCACAGCTATCGTGAAGTCGCCCAACAGCAACACGCTTTCTT 963

1032 TGCAGAGAAGCTTCATCAAGCCATGAAAGTCTTGGAACTCGCCATAGGCATTTGATCAG 1091

964 TGCAGAGAAGCTTCATCAAGCCATGAAAGTCTTGGAACTCGCCATAGGCATTTGATCAG 1023

1092 GATTATGCTTTCCCGGTTCTGAAATTGA 1118

1024 GATTATGG-TTCCCGGTTCTRAAATKRM 1049

RESULT 13

AL551427 1164 bp mRNA linear EST 31-MAY-2003

LOCUS

DEFINITION

AL551427 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA

clone CS0D1064YE07 5-PRIME, mRNA sequence.

ACCESSION

AL551427

VERSION

AL551427.2 GI:31273243

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS

1 (bases 1 to 1164)

Li,W.B., Gruber,C., Jesse,J. and Polayes,D.

TITLE

Full-length cDNA libraries and normalization

JOURNAL

Unpublished (2001)

COMMENT

On Feb 13, 2001 this sequence version replaced gi:12889363.

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 5436.f For

more information about this cluster, see

http://www.genoscope.cns.fr/

cgi-bin/cluster.cgi?seq=CS0D1064AC04QP1&cluster=5436.f. Contact :

Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/ invitrogen Corporation 1600

Faraday Avenue Genoscope sequence ID : CS0D1064AC04QP1.

Location/Qualifiers

1. 1164

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0D1064YE07"

/tissue_type="PLACENTA COT 25-NORMALIZED"

/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT)

FEATURES

source

501 GACATGACGAGGATACATCATTCAGGCTCTTCGAGGATCTGGCCAAAGACATACCTCAG 709


```
841 ACATTAAAGGGTCTACAGAGAGAACTGAAGAGAGATCTGGCCAAAGACATAAACCTCAG 782
710 ACACATCTGAGATTTTCGGAAGCGTTTCTTCTTCTTAAGGGTGACCGATCTGAGG 769
781 ACACATCTGAGATTTTCGGAAGCGTTTCTTCTTCTTAAGGGTGACCGATCTGAGG 722
770 ACTTTGGTGTGAATGAAGACTTGGCTGATTCAGATGCCAGGGCTTGTATGAAGCAGGAG 829
721 ACTTTGGTGTGAATGAAGACTTGGCTGATTCAGATGCCAGGGCTTGTATGAAGCAGGAG 662
830 AAGGAGAAAGGGGACAGAGCTGAACGTTTCAATACCATCCCTTACACAGAGCTATC 889
661 AAGGAGAAAGGGGACAGAGCTGAACGTTTCAATACCATCCCTTACACAGAGCTATC 602
890 CACAACCTTCGAGAGTGTTCAGAAATACACCAAGTACAGTAAAGTACATGAACAAAG 949
601 CACAACCTTCGAGAGTGTTCAGAAATACACCAAGTACAGTAAAGTACATGAACAAAG 542
950 TTTCTGGACCTGGAGTTGAAAGGTGACATTCAGAAATGCGCTCACAGTATCGTGAAGTGG 1009
541 TTTCTGGACCTGGAGTTGAAAGGTGACATTCAGAAATGCGCTCACAGTATCGTGAAGTGG 482
1010 CCACAAGCAACACAGCTTTCTTTTCAGAGAGGCTTTCATCAAGCCATGAAGGTTGTTGAA 1069
481 CCACAAGCAACACAGCTTTCTTTTCAGAGAGGCTTTCATCAAGCCATGAAGGTTGTTGAA 422
1070 CTCGCCATPAAGGCATTCATCAGGATTAATGTTTCCCGTTCTGAAATTCAGATGAATGATA 1129
421 CTCGCCATPAAGGCATTCATCAGGATTAATGTTTCCCGTTCTGAAATTCAGATGAATGATA 362
1130 TCAAGGATTCATCAGAGAGATGATGATATCTCCCTTTCGCAAGCCATCCTCGATGAAA 1189
361 TCAAGGATTCATCAGAGAGATGATGATATCTCCCTTTCGCAAGCCATCCTCGATGAAA 302
1190 CCAAGGAGAGATGAGAAAATCCTGGTGGCTTTTGTGAGAGAACTAAACATTTCCCTT 1249
301 CCAAGGAGAGATGAGAAAATCCTGGTGGCTTTTGTGAGAGAACTAAACATTTCCCTT 242
1250 GATGCTCTCAGCTATGATCAGAGAGCTTTAAATATATATTTTTCATCCTATAAGCTTAAA 1309
241 GATGCTCTCAGCTATGATCAGAGAGCTTTAAATATATATTTTTCATCCTATAAGCTTAAA 182
1310 TAGGAAGATTTCTTCAACAGGATTCAGTGTAGTACTCCTACATGCTGCAAAAATATAGCCT 1369
181 TAGGAAGATTTCTTCAACAGGATTCAGTGTAGTACTCCTACATGCTGCAAAAATATAGCCT 122
1370 TTAATCATTTTATATATTAACCTGTATATAGATAGTCCATTTTAAATATGT 1429
121 TTAATCATTTTATATATTAACCTGTATATAGATAGTCCATTTTAAATATGT 62
1430 TTTCCCAACCAATAAAACCCTATACAGTGTGTTCTAGTAAACATACATAGAGAAAGATGT 1489
61 TTTCCCAACCAATAAAACCCTATACAGTGTGTTCTAGTAAACATACATAGAGAAAGATGT 2
1450 C 1490
1 C 1
```

RESULT 15
AL553095/c
LOCUS 1201 bp mRNA linear EST 31-MAY-2003
DEFINITION AL553095 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA
clone CSODI074YC03 3-PRIME, mRNA sequence.
ACCESSION AL553095
VERSION AL553095.2 GI:31274909
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1201)

AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 15, 2001 this sequence version replaced gi:12892612.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqre@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 5436.f
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CSODI074AB02NP1&cluster=5436.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CSODI074AB02NP1.
Location/Qualifiers
1. 1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODI074YC03"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="First strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

FEATURES
source

ORIGIN

Query Match 62.4%; Score 946.6; DB 9; Length 1201;
Best Local Similarity 96.8%; Pred. NO. 3.2e-241;
Matches 959; Conservative 14; Mismatches 17; Indels 1; Gaps 1;

QY 514 GAGGTGTTTGTAGCTCTGCTAAAACTCCAGGCAATTTGATGCTGATGAACCTCGTCT 573
DB 994 GAGGAGTTGTTTGTAGCTCTGCTAAAACTCCAGGCAATTTGATGCTGATGAACCTCGTCT 935
QY 574 GCCATGAAGGCGCTTGGAACTGATGAAGATCTCTAAATGAGATTTTGGCATCAAGAACT 633
DB 934 GCCATGAAGGCGCTTGGAACTGATGAAGATCTCTAAATGAGATTTTGGCATCAAGAACT 875
QY 634 AACAAGAAATCAGAGACATTAACAGGGTCTACAGAGAGGAACCTGAAGAGAGATCTGGCC 693
DB 874 AACAAGAAATCAGAGACATTAACAGGGTCTACAGAGAGGAACCTGAAGAGAGATCTGGCC 815
QY 694 AAAGACATAAACCTCAGACACATCTGGAGATTTTCGAAACGCTTTTCTTCTTCTTAAG 753
DB 814 AAAGACATAAACCTCAGACACATCTGGAGATTTTCGAAACGCTTTTCTTCTTCTTAAG 755
QY 754 GGTGACCGATCTGAGGACTTTGGTGTGAATGAGACTTTGGCTGATTCAGATGCCAGGGCC 813
DB 754 GGTGACCGATCTGAGGACTTTGGTGTGAATGAGACTTTGGCTGATTCAGATGCCAGGGCC 695
QY 814 TTGTATGAAGCAGGAGAGAGAGGGAAGGGGACAGACGTTAAACGCTGTTCAATACCATCTT 873
DB 694 TTGTATGAAGCAGGAGAGAGAGGAGAGAGGAGAGAGGAGAGAGGAGAGAGGAGAGAG 635
QY 874 ACCACAGAGCTATCCCAACTTCGAGAGTGTTCAGAAATACACCAAGTACAGTAG 933
DB 634 ACCACAGAGCTATCCCAACTTCGAGAGTGTTCAGAAATACACCAAGTACAGTAGTAG 575
QY 934 CATGACATCAACAAAGTTCTGGACCTGGAGTTGAAAGGTGACATTCAGAAATGCTCTACA 993
DB 574 CATGACATCAACAAAGTTCTGGACCTGGAGTTGAAAGGTGACATTCAGAAATGCTCTACA 515
QY 994 GCTATCGTGAAGTGGCCCAAGCAAGCAACGCTTTTCTTTCAGAGAGAGCTTCATCAGGCC 1053
DB 514 GCTATCGTGAAGTGGCCCAAGCAAGCAACGCTTTTCTTTCAGAGAGAGCTTCATCAGGCC 455
QY 1054 ATGAAGAGTGTGGAACTCGCCATAGGCAATGATCAGGATTTATGTTTCCCGTCTGAA 1113
DB 454 ATGAAGAGTGTGGAACTCGCCATAGGCAATGATCAGGATTTATGTTTCCCGTCTGAA 395

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

DM nucleic - nucleic search, using sw model
Run on: March 10, 2004, 08:38:17 ; Search time 6354.47 Seconds
(without alignments)
17011.249 Million cell updates/sec

Title: US-10-084-817-18
Perfect score: 2494
Sequence: 1 gcgcaggatcagcgccgtt.....ccattaaaaaaaaaaaaa 2494

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_man.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2477.8	99.4	2488	9	BC0000013	BC0000013 Homo sapi
2	2457.4	98.5	2549	6	I09499	I09499 Sequence 20
3	2456	98.5	2496	6	AX281786	AX281786 Sequence
4	2446.2	98.1	2499	9	HG1GFBP3M	X64875 H. sapiens m
5	2446.2	98.1	2550	6	AR364291	AR364291 Sequence
6	2445.8	98.1	2497	9	BC018962	BC018962 Homo sapi
7	2430.4	97.4	2474	9	HUMGTBPA	M31159 Human growt
8	2361.8	94.7	2448	6	AX713809	AX713809 Sequence
9	2361.8	94.7	2448	6	AK055632	AK055632 Homo sapi
10	2309.4	92.6	2430	9	AK122873	AK122873 Homo sapi
11	2283.4	91.6	2412	9	AK075483	AK075483 Homo sapi
12	2248.6	90.2	2364	9	AK095907	AK095907 Homo sapi
13	2240.8	89.8	2343	9	AK095869	AK095869 Homo sapi
14	2229.4	89.4	2386	9	AK095408	AK095408 Homo sapi
15	2209.8	88.6	2372	9	AK056274	AK056274 Homo sapi
16	2162.6	86.7	2355	6	AX833836	AX833836 Sequence
17	2162.6	86.7	2355	9	AK095900	AK095900 Homo sapi
18	2156.8	86.5	2312	9	AK091745	AK091745 Homo sapi
19	2152.6	86.3	2343	9	AK095991	AK095991 Homo sapi
20	2143.4	85.9	2191	6	AX014094	AX014094 Sequence
21	2143.4	85.9	2191	6	BD204664	BD204664 Human nuc
22	2101.4	84.3	2327	9	AK127537	AK127537 Homo sapi
23	2099.8	84.2	2327	6	AX747551	AX747551 Sequence
24	2099.8	84.2	2327	9	AK092447	AK092447 Homo sapi
25	2055.4	82.4	2208	9	AK125224	AK125224 Homo sapi
26	2050.8	82.2	2207	9	AK097201	AK097201 Homo sapi
27	2024.8	81.2	2242	9	AK091706	AK091706 Homo sapi
28	1977.6	79.3	2192	6	AX833876	AX833876 Sequence
29	1977.6	79.3	2192	9	AK095974	AK095974 Homo sapi
30	1968	78.9	2141	9	AK123735	AK123735 Homo sapi
31	1943.8	77.9	2173	6	AX833820	AX833820 Sequence
32	1943.8	77.9	2173	9	AK095864	AK095864 Homo sapi
33	1898	76.1	2099	9	AK129951	AK129951 Homo sapi
34	1882.4	75.5	2011	9	AK130884	AK130884 Homo sapi
35	1474.6	59.1	3708	9	HSM807112	EX640961 Homo sapi
36	1471.8	59.0	69887	9	AC091524	AC091524 Homo sapi
37	1460.6	58.6	141539	9	AC146152	AC146152 Pan trogl
38	1460.6	58.6	182079	2	AC146117	AC146117 Pan trogl
39	1459.8	58.5	10884	6	AX409699	AX409699 Sequence
40	1459.8	58.5	10884	9	HUMIBP3	M35878 Homo sapien
41	1426.4	57.2	2438	11	G06425	G06425 human STS W
42	1104.6	44.3	2430	4	AF305199	AF305199 Bos tauru
43	968.4	38.8	2432	10	BC058261	BC058261 Mus muscu
44	914.4	36.7	916	6	AR272361	AR272361 Sequence
45	914.4	36.7	916	6	AR275942	AR275942 Sequence

ALIGNMENTS

RESULT 1
BC0000013
LOCUS
DEFINITION
BC0000013 Homo sapiens insulin-like growth factor binding protein 3, mRNA
(cdna clone MGC:2305 IMAGE:350666), complete cds.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2488)
REFERENCE
AUTHORS
Straussberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenman,C.M., Schuler,G.D.,

Alteschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Seares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,K.D., Mullany,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalius,D.E., Schnerch,A., Schein,J.E., Jones,S.J., and Marra,M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22398257

12477932

2 (bases 1 to 2488)

Strausberg,R.

Direct Submission

Submitted (03-NOV-2000) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

On Aug 20, 2003 this sequence version replaced gi:12652546.

Contact: MGC help desk

Email: cgabbs@mail.nih.gov

Tissue Procurement: DCTD/DTF

cDNA Library Preparation: Rubin Laboratory

cDNA Library Prepared by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Institute for Systems Biology

<http://www.systemsbio.org>

contact: amadan@systemsbiology.org

Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodriguez, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAL Plate: 7 Row: e Column: 4

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 19923110.

Location/Qualifiers

1..2488

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC12305 IMAGE:3506666"

/tissue_type="Kidney, renal cell adenocarcinoma"

/clone_lib="NIH MGC_14"

/lab_host="DH10B-R"

/note="Vector: pOTB7"

1..2488

/gene="IGFBP3"

/note="Synonym: IBP3"

/db_xref="LocustID:3486"

/db_xref="MIM:146732"

99..974

/codon_start=1

/product="insulin-like growth factor binding protein 3"

/protein_id="AAH00013.1"

/db_xref="GI:14652547"

/db_xref="JaccusID:3486"

/translation="MQRARPTLWAAALTLVLRRPVPVARAGSSGGIGPVRCEPCDAPALAQCPAPVACELVREPCGCGCLTALCEQPCGIYTERCGSLRCOPSPDEARPLQALLDGRGLCVNAELVRSLRAYLLPAPAPGNASESEDRSAGSVESPSVSTHVRSDPKPLHLSKIIITKKGHAKDSORYKVDYESQSTDTONFSSKRETEYGPCKRMEPTLNLKELNVLSPSGVHPNCDKKGFKYKKQCRPSKRGKRCWCVDKYQQLPGYT"TKGKEDVHCYSMSQSK"

misc_feature 210..446

/note="IB; Region: Insulin growth factor-binding protein homologues"

/db_xref="CDD:smart00121"

735..953

/note="Thyroglobulin 1; Region: Thyroglobulin type-1 repeat. Thyroglobulin type 1 repeats are thought to be involved in the control of proteolytic degradation. The domain usually contains six conserved cysteines. These form three disulphide bridges. Cysteines 1 pairs with 2, 3 with 4 and 5 with 6"

/db_xref="CDD:pfam00086"

ORIGIN

Query Match 99.4%; Score 2477.8; DB 9; Length 2488;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 2479; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 14 GCGGCTTCCTGCTGGATTCCACAGCTTCGCGCGGTGACTGTGTCGCCCATCCCTGCGGG 73

DB 1 GCGGCTTCCTGCTGGATTCCACAGCTTCGCGCGGTGACTGTGTCGCCCATCCCTGCGGG 60

QY 74 CCCAGCTCCCAAGCAGCGTGCCTGGTTCAGGGGTCTATGCGAGGGGGCGACCCAGCG 133

DB 61 CCCAGCTCCCAAGCAGCGTGCCTGGTTCAGGGGTCTATGCGAGGGGGCGACCCAGCG 120

QY 134 TCTGGGCGCGTGGTGGTACTCTGCTGGTGGTCTCTCGCGGGCGCGCGGTGGCGGGTGG 193

DB 121 TCTGGGCGCGTGGTGGTACTCTGCTGGTGGTCTCTCGCGGGCGCGCGGTGGCGGGTGG 180

QY 194 GCGGAGCTCGGGGGGGCTGGGTTCGGTTCGGTTCGGTTCGGAGCGGTGCGAGCGCGGTGAC 253

DB 181 GCGGAGCTCGGGGGGGCTGGGTTCGGTTCGGTTCGGTTCGGAGCGGTGCGAGCGCGGTGAC 240

QY 254 TGGCCAGTGGCGCGCTCCGCGCGCGGTGTCGGCGGAGCTGTCGGCGAGCGCGGGGTGG 313

DB 241 TGGCCAGTGGCGCGCTCCGCGCGCGGTGTCGGCGGAGCTGTCGGCGAGCGCGGGGTGG 300

QY 314 GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 373

DB 301 GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360

QY 374 GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 433

DB 361 GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420

QY 434 TGGACGGCGCGGGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 493

DB 421 TGGACGGCGCGGGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480

QY 494 TGGACGGCGCGGGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 553

DB 481 TGGACGGCGCGGGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540

QY 554 GTGTGGAGAGCGCGTCCGCTCTCCAGCAGCAGCGGGTGTCTGATCCCAAGTTCACCCCC 613

DB 541 GTGTGGAGAGCGCGTCCGCTCTCCAGCAGCAGCGGGTGTCTGATCCCAAGTTCACCCCC 600

QY 614 TCCATTCAAGATATCATCATCAAGAAAGGCGTCTGTAAGACAGCGCGCTACCAAG 673

DB 601 TCCATTCAAGATATCATCATCAAGAAAGGCGTCTGTAAGACAGCGCGCTACCAAG 660

QY 674 TTGACTACGAGTCTCAGACAGATACCCAGAACTTCTCTCCGAGTCCAGCGGGAGA 733

DB 661 TTGACTACGAGTCTCAGACAGATACCCAGAACTTCTCTCCGAGTCCAGCGGGAGA 720

QY 734 CAGAATATGCTCCCTGCGCTGAGAAATGGAAGACACACTGAATCACCTGAAGTTCCTCA 793

DB 721 CAGAATATGCTCCCTGCGCTGAGAAATGGAAGACACACTGAATCACCTGAAGTTCCTCA 780

QY 794 ATGTGCTGAGTCCCGAGGGGTGTACACATTCCCACTGTGCAAGAGGAGGATTTTAAAG 853

DB 781 ATGTGCTGAGTCCCGAGGGGTGTACACATTCCCACTGTGCAAGAGGAGGATTTTAAAG 840

2y	854	AAAAGCAGTGTCCGCCCTTCCAAAGCGCAGGACGGGCGCTTCTCTGGTGTGGATAAGT	913
2b	841	AAAAGCAGTGTCCGCCCTTCCAAAGCGCAGGACGGGCGCTTCTCTGGTGTGGATAAGT	900
2y	914	ATGGCGACGCTCTCCACGGCTACACCACCACCAAGGGGAGGAGGACGTGCATCTGCATCAGCA	973
2b	901	ATGGCGACGCTCTCCACGGCTACACCACCACCAAGGGGAGGAGGACGTGCATCTGCATCAGCA	960
2y	974	TGCAGACAAAGTAGACGCCCTCGCCGAAGGTTAATGTGGAGCTCAAAATATGCGCTTATTTTG	1033
2b	961	TGCAGACAAAGTAGACGCCCTCGCCGAAGGTTAATGTGGAGCTCAAAATATGCGCTTATTTTG	1020
2y	1034	CACAAAGACGTGCCAAGGACATGACCCAGCAGCTGGCTACAGCCTCGATTTATATTTCTGT	1093
2b	1021	CACAAAGACGTGCCAAGGACATGACCCAGCAGCTGGCTACAGCCTCGATTTATATTTCTGT	1080
2y	1094	TTGTGTGAATCATTTTTTTTTTAAACCAAAGTTTAGAAAAGAGGTTTTTGAAATGCGCTAT	1153
2b	1081	TTGTGTGAATCATTTTTTTTTTAAACCAAAGTTTAGAAAAGAGGTTTTTGAAATGCGCTAT	1140
2y	1154	GGTTCTTTTGAATGGTAAACTTGAGCATCTTTTCACCTTCCAGCTAGTCACGAAGAGCAG	1213
2b	1141	GGTTCTTTTGAATGGTAAACTTGAGCATCTTTTCACCTTCCAGCTAGTCACGAAGAGCAG	1200
2y	1214	TTTGAATTTTCTTGTGCGCTTCCCTATCAAAATATTTACAGAGACTCGAGCACAGCACCCAGAC	1273
2b	1201	TTTGAATTTTCTTGTGCGCTTCCCTATCAAAATATTTACAGAGACTCGAGCACAGCACCCATAC	1260
2y	1274	TTTCATGGCCCGTGGAAATGCTCACCACATGTTTGGTCGAAGCGCGCACCTGACTTGT	1333
2b	1261	TTTCATGGCCCGTGGAAATGCTCACCACATGTTTGGTCGAAGCGCGCACCTGACTTGT	1320
2y	1334	GACTTAGCGGCGCTGTGTGCTATGTAGAGAAACACGCTTACCCCCACCTCCCCGTACAGT	1393
2b	1321	GACTTAGCGGCGCTGTGTGCTATGTAGAGAAACACGCTTACCCCCACCTCCCCGTACAGT	1380
2y	1394	GGCGCAGGCTTTATCGAGAAATAGGAAACCTTTAAACCCCGGTCTACGGACATCCCAA	1453
2b	1381	GGCGCAGGCTTTATCGAGAAATAGGAAACCTTTAAACCCCGGTCTACGGACATCCCAA	1440
2y	1454	CGCATGCTCTGGAGCTCACAGCGCTTCTGTGTGTCTATTTCTGAAAACAGGGCGTGGATC	1513
2b	1441	CGCATGCTCTGGAGCTCACAGCGCTTCTGTGTGTCTATTTCTGAAAACAGGGCGTGGATC	1500
2y	1514	CCTCAACCAAGAGAAATGTTTATGTCTTCAAGTGACCTGTACTGCTTTGGGACTATTGGA	1573
2b	1501	CCTCAACCAAGAGAAATGTTTATGTCTTCAAGTGACCTGTACTGCTTTGGGACTATTGGA	1560
2y	1574	GAAAAAAGGTGGAGTCTACTTGTTTTAAAAAATATGTATCTAAGAAATGTTCTAGGGCAC	1633
2b	1561	GAAAAAAGGTGGAGTCTACTTGTTTTAAAAAATATGTATCTAAGAAATGTTCTAGGGCAC	1620
2y	1634	TTCTGGGAACCTATAAAGGCAAGTATTTTCGGGCCCTTCCTCTTCAGGAACTTCTCTGAAGAC	1693
2b	1621	TTCTGGGAACCTATAAAGGCAAGTATTTTCGGGCCCTTCCTCTTCAGGAACTTCTCTGAAGAC	1680
2y	1694	ATGGCCAGTTCGAAGGCCACAGGATGGCTTTTGTGGGCCCGCTGGGCTAGGAGGACAG	1753
2b	1681	ATGGCCAGTTCGAAGGCCACAGGATGGCTTTTGTGGGCCCGCTGGGCTAGGAGGACAG	1740
2y	1754	AGACACAGGAGAGTCAAGCTCCACATTCAGAGGCATCAACAAGTAATGGCACAAATCTTTC	1813
2b	1741	AGACACAGGAGAGTCAAGCTCCACATTCAGAGGCATCAACAAGTAATGGCACAAATCTTTC	1800
2y	1814	GGATGACTCGAGAAAAATAGTGTGTTTGTAGTTCACAACTCAAGACGAAGCTTATTTCTGA	1873
2b	1801	GGATGACTCGAGAAAAATAGTGTGTTTGTAGTTCACAACTCAAGACGAAGCTTATTTCTGA	1860
2y	1874	GGATAAGCTCTTTAAAGGCAAGGCTTTATTTTCACTCTCATCTTTTGTCCCTTAGCA	1933
2b	1861	GGATAAGCTCTTTAAAGGCAAGGCTTTATTTTCACTCTCATCTTTTGTCCCTTAGCA	1920

Qy	1934	CAATGTAATAAAGAAATAGTAAATATCAAGACAGGAAGGAATGGCTTGCTGGGGAGCCC	1999
Db	1921	CAATGTAAAAAAGAAATAGTAAATATCAGAACAGGAAGGAATGGCTTGCTGGGGAGCCC	1980
Qy	1994	ATCCAGGACACTGGGAGCACATAGAGATTACCCCATGTTTGTGGTGAACCTTAGAGCTCATTTCT	2053
Db	1981	ATCCAGGACACTGGGAGCACATAGAGATTACCCCATGTTTGTGGTGAACCTTAGAGCTCATTTCT	2040
Qy	2054	CATGCTTTTCTTTATTAATTCACACATATATGCAGAGAAGATATGTTCTTGTGTTAAACATTGT	2113
Db	2041	CATGCTTTTCTTTATTAATTCACACATATATGCAGAGAAGATATGTTCTTGTGTTAAACATTGT	2100
Qy	2114	ATCAACATAGCCCCCAATAATAGTAAAGATCTATCTACTAGATAATCTCTAGATGAATGTTAG	2173
Db	2101	ATCAACATAGCCCCCAATAATAGTAAAGATCTATCTACTAGATGAATGTTAG	2160
Qy	2174	AGATGCTATTGTGATAACAACGTGGCATAGCTGAGGAAAGGAGCTCACGCCACAGAGACTG	2233
Db	2161	AGATGCTATTGTGATAACAACGTGGCATAGCTGAGGAAAGGAGCTCACGCCACAGAGACTG	2220
Qy	2234	GGGTGCTCTCCCGGAGGCCAAACCCCAAGAGGTCGCAAAAGTCAGGCTCAGGGAGACTC	2293
Db	2221	GGGTGCTCTCCCGGAGGCCAAACCCCAAGAGGTCGCAAAAGTCAGGCTCAGGGAGACTC	2280
Qy	2294	TGCCCTGCTGCAGACCTCGGTGTGGGACACACGCTGCATAGAGCTCTCTTGAAAAACAGAG	2353
Db	2281	TGCCCTGCTGCAGACCTCGGTGTGGGACACACGCTGCATAGAGCTCTCTTGAAAAACAGAG	2340
Qy	2354	GGGTCTCAAGACACTCTGCCTACCTATTAGCTTTCTTTATTTTAACTTTTTCGGGG	2413
Db	2341	GGGTCTCAAGACACTCTGCCTACCTATTAGCTTTCTTTATTTTAACTTTTTCGGGG	2400
Qy	2414	GAAGAAGTATTTTGAAGAAGTTTCTTGTCAATGATTTTATAAATAGTAATAAAGTTTTT	2473
Db	2401	GAAGAAGTATTTTGAAGAAGTTTCTTGTCAATGATTTTATAAATAGTAATAAAGTTTTT	2460
Qy	2474	ACCATTAAAAAABAAAAA 2494	
Db	2461	ACCATTAAAAAABAAAAA 2481	
RESULT 2			
LOCUS	109499	2549 bp	DNA
DEFINITION	Sequence 20 from Patent WO 8909268.	linear	PAT 02-DEC-1994
ACCESSION	109499		
VERSION	109499.1		
KEYWORDS	GI:587805		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 2549)		
AUTHORS	Baxter,R.C. and Wood,W.I.		
TITLE	PRODUCTION OF INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN		
JOURNAL	Patent: WO 8909268-A 20 05-OCT-1989;		
FEATURES	Location/Qualifiers		
source	1..2549		
	/organism="unknown"		
	/mol_type="unassigned DNA"		
ORIGIN			
Query Match 98.5%; Score 2457.4; DB 6; Length 2549;			
Best Local Similarity 99.9%; Pred. No. 0;			
Matches 2480; Conservative 0; Mismatches 1; Indels 2; Gaps 2;			
Qy	12	CAGCCGGTCTCCTCGCTGGATTCCACAGCTTCGGCGCGTGTACTGTGCGCCCACTCCCTGCG	71
Db	10	CAGCCGGTCTCCTCGCTGGATTCCACAGCTTCGGCGCGTGTACTGTGCGCCCACTCCCTGCG	69
Qy	72	CGCCAGCTTCGCCAAGCAGCGTGCCTCGGTGTGACGCGCTCATGACGGGGCGGACCCAC	131
Db	70	CGCCAGCTTCGCCAAGCAGCGTGCCTCGGTGTGACGCGCTCATGACGGGGCGGACCCAC	129

QY 132 GCTGTGGCCGCTGCGCTGAGATCTGCTGGTGGTCTGCTCCGCGGGCCGCGGTGGCGCGGC 191
 Db 130 GCTGTGGCCGCTGCGCTGAGATCTGCTGGTGGTCTGCTCCGCGGGCCGCGGTGGCGCGGC 189
 QY 192 TGGCGCAGAGCTCGGGGGCTTGGGTCCGCTGCTGGCTGCGAGCCGTCGACGCGCGTGC 251
 Db 190 TGGCGCAGAGCTCGGGGGCTTGGGTCCGCTGCTGGCTGCGAGCCGTCGACGCGCGTGC 249
 QY 252 ACTGGCCAGTGGCGGCTTCCGCGCGCGCTGCGCGGAGAGTGTGGCGGAGCGCGGCTG 311
 Db 250 ACTGGCCAGTGGCGGCTTCCGCGCGCGCTGCGCGGAGAGTGTGGCGGAGCGCGGCTG 309
 QY 312 CGGCTGCTGCTGAGCGTGGCACTGAGGAGGCGCGAGCCGTCGCGGCATCTACACCGAGCG 371
 Db 310 CGGCTGCTGCTGAGCGTGGCACTGAGGAGGCGCGAGCCGTCGCGGCATCTACACCGAGCG 369
 QY 372 CTGTGGCTCCGCGCTTCCGCTGCGCGCGCTGCGCGGAGCGCGAGCGCGCTGCGAGGCGCT 431
 Db 370 CTGTGGCTCCGCGCTTCCGCTGCGCGCGCTGCGCGGAGCGCGAGCGCGCGCTGCGAGGCGCT 429
 QY 432 GCTGAGCGCGCGGCTTCCGCTGCGCGCTGAGTGGCTGCGCGCTGCGCGCTGCGCGCTTACCT 491
 Db 430 GCTGAGCGCGCGGCTTCCGCTGCGCGCTGAGTGGCTGCGCGCTGCGCGCTTACCT 489
 QY 492 GCTGCGAGCGCGCGCTTCCGAGGAAATGCTAGTGGTGGTGGGAGGAGCGCGAGCGCGG 551
 Db 490 GCTGCGAGCGCGCGCTTCCGAGGAAATGCTAGTGGTGGTGGGAGGAGCGCGAGCGCGG 549
 QY 552 CAGTGTGAGAGCGCGCTTCCGCTGCGCGAGCGCGAGCGCGCTGCGCGCTGCGCGCTTACCT 611
 Db 550 CAGTGTGAGAGCGCGCTTCCGCTGCGCGAGCGCGAGCGCGCTGCGCGCTTACCT 609
 QY 612 CCTCCATTCAAGATAATCATCATCAAGAAAGGCGATGCTAAAGACAGCGCGCTACAA 671
 Db 610 CCTCCATTCAAGATAATCATCATCAAGAAAGGCGATGCTAAAGACAGCGCGCTACAA 669
 QY 672 AGTTGACTAGAGTCTAGAGCAGATACCGAGAACTTCTTCTCCGAGTCCAGAGCGGA 731
 Db 670 AGTTGACTAGAGTCTAGAGCAGATACCGAGAACTTCTTCTCCGAGTCCAGAGCGGA 729
 QY 732 GACAGAAATGCTTCCGCTGCGTGGAGAAATGGAAGACACACTGAATCACCTGGAAGTTTCT 791
 Db 730 GACAGAAATGCTTCCGCTGCGTGGAGAAATGGAAGACACACTGAATCACCTGGAAGTTTCT 789
 QY 792 CAAATGTGTAGTCCAGGGGTGTACATTCCTCCAACTGTGACAGAGAGGATTTATAA 851
 Db 790 CAAATGTGTAGTCCAGGGGTGTACATTCCTCCAACTGTGACAGAGAGGATTTATAA 849
 QY 852 GAAAAGCAGTGTGCTTCCAAAGGAGGAGGAGCGGGCTTCTGCTGCTGTGGATAA 911
 Db 850 GAAAAGCAGTGTGCTTCCAAAGGAGGAGGAGCGGGCTTCTGCTGCTGTGGATAA 909
 QY 912 GTATGGGAGCTTCCAGGCTACACCAAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 971
 Db 910 GTATGGGAGCTTCCAGGCTACACCAAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 969
 QY 972 CATGAGAGCAGTACAGCTTCCGCGAGGATGAGTGGAGCTCAATATGCTTATTT 1031
 Db 970 CATGAGAGCAGTACAGCTTCCGCGAGGATGAGTGGAGCTCAATATGCTTATTT 1029
 QY 1032 TGCAAAAAGACTGCGCAAGGACATGACGAGCTGGCTCAGAGCTCGATTTATATTTCT 1091
 Db 1030 TGCAAAAAGACTGCGCAAGGACATGACGAGCTGGCTCAGAGCTCGATTTATATTTCT 1089
 QY 1092 GTTGTGGTGAATGATTTTTTTTAAACCAAGTTTAAAGAGGTTTTTAAATGCTT 1151
 Db 1090 GTTGTGGTGAATGATTTTTTTTAAACCAAGTTTAAAGAGGTTTTTAAATGCTT 1148
 QY 1152 ATGGTTTCTTGAATGGTAAATTTGAGCATCTTTTCACTTCCAGTGTGAGCAAGAGC 1211
 Db 1149 ATGGTTTCTTGAATGGTAAATTTGAGCATCTTTTCACTTCCAGTGTGAGCAAGAGC 1208
 QY 1212 AGTTTGAATTTCTTGTGCTTCTTATCAAAATATTTAGAGACTGCGAGCAGCAGCCAG 1271

Db 1209 AGTTTGAATTTCTTGTGCTTCTTATCAAAATATTTAGAGACTGAGCAGCAGCCAG 1269
 QY 1272 ACTTCATGCGCGCGTGGAAATGCTCACCACATGTTGGTGAAGCGGCCAGCAGCTACCTTT 1331
 Db 1269 ACTTCATGCGCGCGTGGAAATGCTCACCACATGTTGGTGAAGCGGCCAGCAGCTACCTTT 1328
 QY 1332 GTGACTTAGCGCGCTGTTGGCTATGTAGAGAACGCTTCAACCCGCTCATCCCGTACA 1391
 Db 1329 GTGACTTAGCGCGCTGTTGGCTATGTAGAGAACGCTTCAACCCGCTCATCCCGTACA 1388
 QY 1392 GTGCGCAGCAGGCTTATCGAGAAATAGGAAACCTTTAAACCCGCTCATCCCGGATCCC 1451
 Db 1389 GTGCGCAGCAGGCTTATCGAGAAATAGGAAACCTTTAAACCCGCTCATCCCGGATCCC 1448
 QY 1452 AAGCATGCTTCTGAGCTCAGCGCTTCTGTGTGTCTTCAAGTGAACCTGCTTGAAGAGGCTGGA 1511
 Db 1449 AAGCATGCTTCTGAGCTCAGCGCTTCTGTGTGTCTTCAAGTGAACCTGCTTGAAGAGGCTGGA 1508
 QY 1512 TCCCTCAACCAAGAAAGATGTTTATGCTTCAAGTGAACCTGCTTGAAGAGGCTGGA 1571
 Db 1509 TCCCTCAACCAAGAAAGATGTTTATGCTTCAAGTGAACCTGCTTGAAGAGGCTGGA 1568
 QY 1572 GAGAAATAGTGGAGTCTTACTTGTTTTAAATAATGATCTTAAGAAATGTTCTAGGCG 1631
 Db 1569 GAGAAATAGTGGAGTCTTACTTGTTTTAAATAATGATCTTAAGAAATGTTCTAGGCG 1628
 QY 1632 ACTCTGGGAACCTATAAAGGAGGATTTTCGGGCCCTCTCTTCAAGAAATCTTCTGAAG 1691
 Db 1629 ACTCTGGGAACCTATAAAGGAGGATTTTCGGGCCCTCTCTTCAAGAAATCTTCTGAAG 1688
 QY 1692 ACATGSCCAGTCTGAAGGCCAGGATGCTTGTGCGGCCCTGCTGGGTAGGAGGAGC 1751
 Db 1689 ACATGSCCAGTCTGAAGGCCAGGATGCTTGTGCGGCCCTGCTGGGTAGGAGGAGC 1748
 QY 1752 AGAGAGCAGGAGGAGTCTGAGCTCCACATTCAGAGGCTCAAGTAAATGCGACAAATCT 1811
 Db 1749 AGAGAGCAGGAGGAGTCTGAGCTCCACATTCAGAGGCTCAAGTAAATGCGACAAATCT 1807
 QY 1812 TCGGATGACTGAGAAATAGTGTGTTTGTAGTTCACAACTCAAGAGGAGGATGCTTCT 1871
 Db 1808 TCGGATGACTGAGAAATAGTGTGTTTGTAGTTCACAACTCAAGAGGAGGATGCTTCT 1867
 QY 1872 GAGGATAAGCTCTTAAAGGCAAGCTTATTTTATCTCTCATCTTGTCTCTCTAG 1931
 Db 1868 GAGGATAAGCTCTTAAAGGCAAGCTTATTTTATCTCTCATCTTGTCTCTCTAG 1927
 QY 1932 CACAATGTAAGAAATAGTAAATATCAGAAAGGAGGAGGATGCTTCTTGGGAGC 1991
 Db 1928 CACAATGTAAGAAATAGTAAATATCAGAAAGGAGGAGGATGCTTCTTGGGAGC 1987
 QY 1992 CMTCCAGGAGCAGTGGAGCAGATAGAGATTCACCATGTTTGTGTAAGTGAAGTCAAT 2051
 Db 1988 CMTCCAGGAGCAGTGGAGCAGATAGAGATTCACCATGTTTGTGTAAGTGAAGTCAAT 2047
 QY 2052 CTATGCTTCTTCTTATAATTCACATATATGAGAGGAGGATGCTTCTTGAAGTCAAT 2111
 Db 2048 CTATGCTTCTTCTTATAATTCACATATATGAGAGGAGGATGCTTCTTGAAGTCAAT 2107
 QY 2112 GTATACAAATAGCCCAATATAGTAAATCTTATAGTAAATCTTAGATGAATGTT 2171
 Db 2108 GTATACAAATAGCCCAATATAGTAAATCTTATAGTAAATCTTAGATGAATGTT 2167
 QY 2172 AGAGATGCTTATGATACAACTGTGCGCATGCTGAGGAAAGGAGCTCACGCCAGAGAGC 2231
 Db 2168 AGAGATGCTTATGATACAACTGTGCGCATGCTGAGGAAAGGAGCTCACGCCAGAGAGC 2227
 QY 2232 TGGGCTGCTCTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2291
 Db 2228 TGGGCTGCTCTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2287
 QY 2292 TCTGCGCTCTGAGAGCTGCGGTGGAGCAGCAGCTGATAGAGCTCTCTTGAAGAGCAG 2351

```
2288 TCTGCCCTGCTGAGACCTCGGTGTGGACACACGCTGCATAGAGCTCTCTCTGAAACAG 2347
2352 AGGGGCTCAAGACATCTGCTACCTACCTATTAGCTTTCTTTCTTTTAACTTTTGGG 2411
2348 AGGGGCTCAAGACATCTGCTACCTACCTATTAGCTTTCTTTCTTTTAACTTTTGGG 2407
2412 GGGAAAAGTATTTTGGAAAGTTTGTCTGTCATGATTTATTAATAGTAAATAGTTT 2471
2408 GGGAAAAGTATTTTGGAAAGTTTGTCTGTCATGATTTATTAATAGTAAATAGTTT 2467
2472 TTACCATTAATAAAAAAAAAAAAA 2494
2468 TTACCATTAATAAAAAAAAAAAAA 2490

RESULT 3
LOCUS AX281786
DEFINITION Sequence 195 from Patent WO0177389.
ACCESSION AX281786
VERSION AX281786.1 GI:16609037
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Shiffman,D., Sonogyi,R., Lawn,R., Seilhamer,J.J., Porter,G.J.,
Mukita,T. and Tai,J.
TITLE Genes expressed in foam cell differentiation
JOURNAL Patent: WO 0177389-A 195 18-OCT-2001;
Incyte Genomics, Inc. (US)
FEATURES
source 1..2496
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 253550.14"

ORIGIN
Query Match 98.5%; Score 2456; DB 6; Length 2496;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2478; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

2y 1 GCGCTGAGGATCAGCCGCTTCTGCTCGTGGATTCACAGCTTCGCGCGTGTACTGCGC 60
Db 18 GCGCTGAGGATCAGCCGCTTCTGCTCGTGGATTCACAGCTTCGCGCGTGTACTGCGC 77
2y 61 CCATCCCTGCGCGCCAGCCCTGCCAA-CCAGCGTCCCGTTCAGGGCGTCAATGAGCG 119
Db 78 CCATCCCTGCGCGCCAGCCCTGCCAAAGCGTCCCGTTCAGGGCGTCAATGAGCG 137
2y 120 GCGCGGACCAACGCTCTGGCGCTGCGCTGACTCTGCTGCTGCTGCTGCTGCTGCGCGCGCC 179
Db 138 GCGCGGACCAACGCTCTGGCGCTGCGCTGACTCTGCTGCTGCTGCTGCTGCTGCGCGCGCC 197
2y 180 GGTGGCGGCGCTGCGCGAGCTGCGGGGCTGGGTCCGCTGCTGCTGCTGCTGCGCGCGCTG 239
Db 198 GGTGGCGGCGCTGCGCGAGCTGCGGGGCTGGGTCCGCTGCTGCTGCTGCTGCGCGCGCTG 257
2y 240 CGAGCGCGCTGCACTGGCCAGTGGCGGCTTCGCGCGCTGCTGCTGCTGCTGCTGCTGCGCG 299
Db 258 CGAGCGCGCTGCACTGGCCAGTGGCGGCTTCGCGCGCTGCTGCTGCTGCTGCTGCTGCGCG 317
2y 300 CGAGCGGCGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGCGCT 359
Db 318 CGAGCGGCGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGCGCT 377
2y 360 CTACACCGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGCGCGCT 419
Db 378 CTACACCGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGCGCGCT 437
2y 420 GCTGAGGCGCTGCTGAGCGCGCGGCTCTGCGCTCAACGCTAGTCCGCTCAGCGCGCT 479
```

```
Db 438 GCTGAGGCGCTGCTGAGCGCGCGGCTCTGGTCAACGCTAGTCCGCTCAGCGCGCT 497
2y 480 GCGGCGCTTACCTGCTGCGAGCGCGCCAGCTCCAGGAAATGCTAGTGTGAGGAGGA 539
Db 498 GCGGCGCTTACCTGCTGCGAGCGCGCCAGCTCCAGGAAATGCTAGTGTGAGGAGGA 557
2y 540 CCGAGCGCGCGAGTGTGGAGCGCGCTCCGCTCCAGCAGCGCCGCGTGTCTGATCC 599
Db 558 CCGAGCGCGCGAGTGTGGAGCGCGCTCCGCTCCAGCAGCGCCGCGTGTCTGATCC 617
2y 600 CAAAGTTCACCCCTCCATTCAAAGATAATCATCATCAAGAAAGGGCATGCTAAAGACAG 659
Db 618 CAAAGTTCACCCCTCCATTCAAAGATAATCATCATCAAGAAAGGGCATGCTAAAGACAG 677
2y 660 CGAGGCTACAAAGTTGACTACGAGTCTCAGAGCAGACATACCCAGACTTCTCTCCGA 719
Db 678 CCAGCGCTTACAAAGTTGACTACGAGTCTCAGAGCAGACATACCCAGACTTCTCTCCGA 737
2y 720 GTCCAAAGCGGAGACAGAATATGCTCCCTGCGGTAGAGAAATGGAAGACACACTGAATCA 779
Db 738 GTCCAAAGCGGAGACAGAATATGCTCCCTGCGGTAGAGAAATGGAAGACACACTGAATCA 797
2y 780 CCGTAAGTTCCTCAATGTGCTGAGTCCAGGGGTGTACACATTCCTCAACTGTGACAGAA 839
Db 798 CCTGAAGTTCCTCAATGTGCTGAGTCCAGGGGTGTACACATTCCTCAACTGTGACAGAA 857
2y 840 GGGATTTTATAAGAAAGACAGTGTGCGCCCTTCCAAAGGAGGAGCGGGCTTCTGCTG 899
Db 858 GGGATTTTATAAGAAAGACAGTGTGCGCCCTTCCAAAGGAGGAGCGGGCTTCTGCTG 917
2y 900 GTGTGTGATTAAGTATGGGCGAGCTCTCCAGGCTACACCAAGGAGGAGGAGGAGCT 959
Db 918 GTGTGTGATTAAGTATGGGCGAGCTCTCCAGGCTACACCAAGGAGGAGGAGGAGCT 977
2y 960 GCACCTGTACAGCATGACAGCAAGTAGAGCCCTGCGCAAGGTTAATGTGGAGCTCAAA 1019
Db 978 GCACCTGTACAGCATGACAGCAAGTAGAGCCCTGCGCAAGGTTAATGTGGAGCTCAAA 1037
2y 1020 TATGCTTTATTTGCAAAAAGATGCGCAAGGACATGACAGCGAGCTGGCTACAGCTCG 1079
Db 1038 TATGCTTTATTTGCAAAAAGATGCGCAAGGACATGACAGCGAGCTGGCTACAGCTCG 1097
2y 1080 ATTATATTTCTGTTTGTGGTGAATGTTTTTTTAAACCAAGTTTAAAGAGAGTT 1139
Db 1098 ATTATATTTCTGTTTGTGGTGAATGTTTTTTTAAACCAAGTTTAAAGAGAGTT 1156
2y 1140 TTTGAAATGCCCTATGTTTCTTTGAATGGTAAACTTGAGCATCTTTTCACTTCCAGTAG 1199
Db 1157 TTTGAAATGCCCTATGTTTCTTTGAATGGTAAACTTGAGCATCTTTTCACTTCCAGTAG 1216
2y 1200 TCAGCAAAAGAGCGAGTTGAATTTTCTTGTGCTTCTTATCAAAATATTTCAGAGACTCGAG 1259
Db 1217 TCAGCAAAAGAGCGAGTTGAATTTTCTTGTGCTTCTTATCAAAATATTTCAGAGACTCGAG 1276
2y 1260 CACAGCACCCAGACTTCATGCGCGCGTGGAAATGCTCACCATGTTGGTCCGAAGCGCGC 1319
Db 1277 CACAGCACCCAGACTTCATGCGCGCGTGGAAATGCTCACCATGTTGGTCCGAAGCGCGC 1336
2y 1320 ACCACTGACTTTGTGACTTAGCGGCTGTGTTGCTATGTAGAGAACAGCTTCAACCCC 1379
Db 1337 ACCACTGACTTTGTGACTTAGCGGCTGTGTTGCTATGTAGAGAACAGCTTCAACCCC 1396
2y 1380 ACTCCCGGTACAGTCCGACAGGCTTTATCGAGAAATAGGAAACCTTTAAACCCCGGTCA 1439
Db 1397 ACTCCCGGTACAGTCCGACAGGCTTTATCGAGAAATAGGAAACCTTTAAACCCCGGTCA 1456
2y 1440 TCGGAGCATCCCAAGCATGCTCTCGAGCTCAAGCCCTCTGTGGTGTCTATTCTGAAA 1499
Db 1457 TCGGAGCATCCCAAGCATGCTCTCGAGCTCAAGCCCTCTGTGGTGTCTATTCTGAAA 1516
2y 1500 CAAAGGCGTGGATCCCTCAACCAAGAAATGTTTATGTTTCAAGTGACCTGACTGCT 1559
```


Db	1517	CAAGGCGTGGATCCTCTCAACCAAGAAAGTGTTTATGTCTTCAAGTGAACCTGTACTGCT	1576
Qy	1560	TGGGGACTATTGGAGAAAAATAAGGTGGAGTCTTACTTGTTTAAAAAATATGTATCTAAAGA	1619
Db	1577	TGGGGACTATTGGAGAAAAATAAGGTGGAGTCTTACTTGTTTAAAAAATATGTATCTAAAGA	1636
Qy	1620	ATGTTCTAGGGCACTCTGGGAACTATAAGGACAGTATTTTGGGGCCCTCTCTTTCAGGA	1679
Db	1637	ATGTTCTAGGGCACTCTGGGAACTATAAGGACAGTATTTTGGGGCCCTCTCTTTCAGGA	1696
Qy	1680	ATCTTCTCTGAAGACATGGCCCACTCGAAGGCCCCAGGATGGCTTTTGTCTCGGGCCCCGTGG	1739
Db	1697	ATCTTCTCTGAAGACATGGCCCACTCGAAGGCCCCAGGATGGCTTTTGTCTCGGGCCCCGTGG	1756
Qy	1740	GGTAGGAGGACACAGAGACACAGGAGAGTCAGCTCTCCACATTCAGAGGCATCACAGTAA	1799
Db	1757	GGTAGGAGGACACAGAGACACAGGAGAGTCAGCTCTCCACATTCAGAGGCATCACAGTAA	1816
Qy	1800	TGGCACAAATCTTTCGGATGACTCGAGAAAATAGTGTTTTGTAGTTTCAACAACCTCAAGACG	1859
Db	1817	TGGCACAAATCTTTCGGATGACTCGAGAAAATAGTGTTTTGTAGTTTCAACAACCTCAAGACG	1876
Qy	1860	AAGCTTATTCTGAGGNTAAGCTCTTTAAAGGCAAGCTTTATTTTCACTCTCTCATCTTT	1919
Db	1877	AAGCTTATTCTGAGGNTAAGCTCTTTAAAGGCAAGCTTTATTTTCACTCTCTCATCTTT	1936
Qy	1920	TGTCCTCTTTAGCACAAATGTAAAAAGAAATAGTAATATCAGAACAGGAAGGAGGAATGGC	1979
Db	1937	TGTCCTCTTTAGCACAAATGTAAAAAGAAATAGTAATATCAGAACAGGAAGGAGGAATGGC	1996
Qy	1980	TTGCTGGGGAGCCCATCCAGGACACTGGGAGCACATAGAGATTCACCCATGTTTGTGAA	2039
Db	1997	TTGCTGGGGAGCCCATCCAGGACACTGGGAGCACATAGAGATTCACCCATGTTTGTGAA	2056
Qy	2040	CTTAGAGTCAATCTCATGCTTTCTTTTATAATTCACACATATATGCAGAGAAGATATGTT	2099
Db	2057	CTTAGAGTCAATCTCATGCTTTCTTTTATAATTCACACATATATGCAGAGAAGATATGTT	2116
Qy	2100	CTTGTTAAATTTGATACACATAGCCCCAAATATAGTAAGTCTTATCTAGATAATCCT	2159
Db	2117	CTTGTTAAATTTGATACACATAGCCCCAAATATAGTAAGTCTTATCTAGATAATCCT	2176
Qy	2160	AGATGAATTTGATAGATGCTATTGTATGATCAACCTGTGGCCATGACTGAGGAAGGAGCTC	2219
Db	2177	AGATGAATTTGATAGATGCTATTGTATGATCAACCTGTGGCCATGACTGAGGAAGGAGCTC	2236
Qy	2220	ACGCCACAGACTGGGTGTCTCTCCGGAGGCCAAACCCAAAGAGTCTGGCAAAAGTCAG	2279
Db	2237	ACGCCACAGACTGGGTGTCTCTCCGGAGGCCAAACCCAAAGAGTCTGGCAAAAGTCAG	2296
Qy	2280	GCTCAGGAGACTCTGCCCTGCTGCAGACTCTGGGTGTGGACACACGCTGCATAGAGCTCT	2339
Db	2297	GCTCAGGAGACTCTGCCCTGCTGCAGACTCTGGGTGTGGACACACGCTGCATAGAGCTCT	2356
Qy	2340	CCTTGAACACAGAGGGTCTCAAGACATCTTGCCATCTATAGCTTTTCTTTATTTTTT	2399
Db	2357	CCTTGAACACAGAGGGTCTCAAGACATCTTGCCATCTATAGCTTTTCTTTATTTTTT	2416
Qy	2400	TAACTTTTGGGGGAAAAGTATTTTGGAGAAGTTGTTCTTGCAATGTATTTATAAATAG	2459
Db	2417	TAACTTTTGGGGGAAAAGTATTTTGGAGAAGTTGTTCTTGCAATGTATTTATAAATAG	2476
Qy	2460	TAAATAAAGTTTTTACCATT	2479
Db	2477	TAAATAAAGTTTTTACCATT	2496

insulin-like growth factor-binding protein 3.
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 2499)
 Thweatt,R., Fleischmann,R. and Goldstein,S.
 Analysis of the primary structure of insulin-like growth factor
 binding protein-3 cDNA from Werner syndrome fibroblasts
 DNA Seq. 4 (1), 43-46 (1993)
 7508771
 2 (bases 1 to 2499)
 Thweatt,R.
 Direct Submission
 Submitted (09-MAR-1992) R. Thweatt, Univ Arkansas for Med Sciences,
 John L McClellan VA, 4300 West 7th Street/Res 151, Little Rock, AR
 72205, USA
 Location/Qualifiers
 1..2499
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /isolate="Werners syndrome subject TO"
 /db_xref="taxon:9606"
 /cell_line="WS8"
 /cell_type="fibroblast"
 /tissue_type="skin"
 /clone_lib="Werners syndrome fibroblast cDNA"
 /dev_stage="adult"
 1..2499
 /gene="IGFBP-3"
 133..1009
 /gene="IGFBP-3"
 /codon_start=1
 /product="insulin-like growth factor binding protein 3"
 /protein_id="CAA46087.1"
 /db_xref="GI:398164"
 /db_xref="GOA:P17936"
 /db_xref="SWISS-PROT:P17936"
 /translation="MQRASPTLWAAALTLLVLRGPPVAGASGGSLGFPWRCEPCD
 ARALACAPGPAVCALRVPEFGCCCLTCAISGQPCGIYTERCGSLRCQPSDEAR
 PQALADGGKCVNAAVSRRLAYLPAPAPNASESDRSAGSVESVPSASTHVA
 SPQFPLHLSKIIIKKGHAKQSQRYKVDSQSTPNFSSKRETEIGPQREHME
 DTNLHLKFLNVLSPRGVHPNCDDKKQFKYKKRGRFCWCVDKYGQPLPGVYT
 TKGKEDVHCYVSMQSK"
 133..213
 /gene="IGFBP-3"
 2482..2486
 /gene="IGFBP-3"
 sig_peptide
 polyA_signal
 ORIGIN
 Query Match 98.1%; Score 2446.2; DB 9; Length 2499;
 Best Local Similarity 99.4%; Pred. No. 0;
 Matches 2465; Conservative 0; Mismatches 13; Indels 1; Gaps 1/
 Qy 1 GCGCTGAGGATCAGCGCTTCCTGCTGGATTCCACAGCTCGCCCGGTTCGAGCGCTCATCGACGG 60
 Db 22 GCGCTGAGGATCAGCGCTTCCTGCTGGATTCCACAGCTTCGCGCGCTGACTGTGCTGCC 81
 Qy 61 CCATCCCTGCGGCCAGCGCTGCCAAGCAGCGTCCCCCGGTTCGAGCGCTCATCGACGG 120
 Db 82 CCATCCCTGCGGCCAGCGCTGCCAAGCAGCGTCCCCCGGTTCGAGCGCTCATCGACGG 141
 Qy 121 GCGGGAACCAAGCTCTGGGGCGCTGCCTGACTGCTGCTGCTGCTCCGGGGACCGCGC 180
 Db 142 GCGGGAACCAAGCTCTGGGGCGCTGCCTGACTGCTGCTGCTGCTCCGGGGACCGCGC 201
 Qy 181 GTGGCGGGGCTGGCGGAGCTCGGGGGGGCTTGGGTCCCGTGGTGGCTGCGAGCCGGTGC 240
 Db 202 GTGGCGGGGCTGGCGGAGCTCGGGGGGCTTGGGTCCCGTGGTGGCTGCGAGCCGGTGC 261
 Qy 241 GACCGCGTGCACCTGGGCCAGTGCAGCGCTCCCGCCCGCGCTGTGTCGGGGAGCTGGTGGC 300


```
1752 AGAGAGACGGGAGAGTCCAGCTCCACATTCAGAGGCATCAAGTAATGGCACAAATCT 1811
b 1749 AGAGAGAC-GGGAGAGTCCAGCTCCACATTCAGAGGCATCAAGTAATGGCACAAATCT 1807
y 1812 TCGGATGACTGCAGAAATAGTGTGTTGTTAGTTCACAACTCAAGCAGAGCTTATTTCT 1871
b 1808 TCGGATGACTGCAGAAATAGTGTGTTGTTAGTTCACAACTCAAGCAGAGCTTATTTCT 1867
y 1872 GAGGATAAGCTCTTTAAAGGCAAAAGCTTTATTTTCATCTCTCATCTTTTGCTCTCTTAG 1931
b 1868 GAGGATAAGCTCTTTAAAGGCAAAAGCTTTATTTTCATCTCTCATCTTTTGCTCTCTTAG 1927
y 1932 CACAATGTAATAAAGATAGTATATATACAGACAGGAGGAATGCTTCTGCTGGGAGC 1991
b 1928 CACAATGTAATAAAGATAGTATATATACAGACAGGAGGAATGCTTCTGCTGGGAGC 1987
y 1992 CCATCCAGACACTGGGAGCACATAGAGATTCACCCATGTTTGTGAACTTAGAGTCATT 2051
b 1988 CCATCCAGACACTGGGAGCACATAGAGATTCACCCATGTTTGTGAACTTAGAGTCATT 2047
y 2052 CTCATGCTTTCTTTTATATATACATATATGCAGAGAGATATGTTCTGTTTAAACATT 2111
b 2048 CTCATGCTTTCTTTTATATATACATATATGCAGAGAGATATGTTCTGTTTAAACATT 2107
y 2112 GTATACACATACCCCAATATAGTAAGATCTATCTAGATAATCCTAGATGAATGTT 2171
b 2108 GTATACACATACCCCAATATAGTAAGATCTATCTAGATAATCCTAGATGAATGTT 2167
y 2172 AGAGATGCTATTTGATACAACTGTGCGCATGACTGAGGAAAGAGCTCAGCCGACAGAC 2231
b 2168 AGAGATGCTATATGATACAACTGTGCGCATGACTGAGGAAAGAGCTCAGCCGACAGAC 2227
y 2232 TGGGCTGCTCTCCGAGGAGCCAAACCAAGAGGCTCTGCAAAAGTCAGGCTCAGGAGAC 2291
b 2228 TGGGCTGCTCTCCGAGGAGCCAAACCAAGAGGCTCTGCAAAAGTCAGGCTCAGGAGAC 2287
y 2292 TCTGSCCTGTGCAGACCTCGGTGTGCACACACAGCTGCATAGAGCTCTCCTTGAAACAG 2351
b 2288 TCTGSCCTGTGCAGACCTCGGTGTGCACACACAGCTGCATAGAGCTCTCCTTGAAACAG 2347
y 2352 AGGGTCTCAGACATCTGCTACCTATAGCTTCTTTTATTTTAACTTTTGGG 2411
b 2348 AGGGTCTCAGACATCTGCTACCTATAGCTTCTTTTATTTTAACTTTTGGG 2407
y 2412 GGGAAAAGTATTTTTCAGAGATTTGCTTGCAATGTATTATAAATAGTAATAAGTTT 2471
b 2408 GGGAAAAGTATTTTTCAGAGATTTGCTTGCAATGTATTATAAATAGTAATAAGTTT 2467
y 2472 TTACCAATATAAAAAAAAAAAAAA 2494
b 2468 TTACCAATATAAAAAAAAAAAAAA 2490
RESULT 6
LOCUS BC018962
DEFINITION Homo sapiens insulin-like growth factor binding protein 3, mRNA
ACCESSION BC018962
VERSION BC018962.2
KEYWORDS G1:33879043
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2497)
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Datchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
```

```
Carinci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettelman, M., Madan, A.C., Rodrigues, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalski, U., Smallus, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
23388257
12477932
2 (bases 1 to 2497)
Strausberg, R.
Direct Submission
Submitted (07-DEC-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
On Aug 19, 2003 this sequence version replaced gi:17511987.
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc.mgc@nih.gov
Akhter, N., Ayala, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghini, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
Maduro, Q.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 27 Row: c Column: 22.
```

FEATURES

```
1. .2497
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:19695 IMAGE:3506929"
/tissue_type="Kidney, renal cell adenocarcinoma"
/clone_lib="NIH MGC 14"
/lab_host="DH10B-R"
/notes="Vector: pOTB7"
1. .2497
/gene="IGFBP3"
/notes="synonym: IBP3"
/db_xref="LocusID:3486"
/db_xref="MIM:146732"
79. - 954
/codon_start=1
/product="insulin-like growth factor binding protein 3"
/protein_id="AAH18962.1"
/db_xref="GI:17511988"
/translaton="MORAPTLMAAATLLVLRGPPVAPAGASSGSLGFPVPCPCD
ARALAQCAPPPVACVRLVRLVLLPAPGNASSEDSRGSVESVSTHVR
PQALLDGRGICVNAVSRLVRLVLLPAPGNASSEDSRGSVESVSTHVR
SDPHEPLHLSKILTIKKGAKDORYKDYVESQSTDTQNFSESKRETEYGPCKREME
DTLNLHLKFLNLSRPGVHPCNDKXGFKKKQCPKSKRGKGCWCVCVDYQPLPGYT
```

gene

CDS

Query Match	98.1%	Score 2445.8	DB 9	Length 2497
Best Local Similarity	99.9%	Pred. No. 0		
Matches 2458	Conservative	0	Mismatches	2
			Indels	1
			Gaps	1
QY	34	CACAGCTTCGCGCGGTGTACTGTGCGCCCATCCCTGCGCCCGCCAGCGCTGCCAAGCAGCGT	93	
DB	1	CACAGCTTCGCGCGGTGTACTGTGCGCCCATCCCTGCGCGCCCGCCAGCGCTGCCAAGCAGCGT	60	
QY	94	GCCCCGGTTGCAGGCGGTTCATGCAAGCGGCGCGACCCACAGCTCTCTGGGCGGTGCGCTGCACT	153	
DB	61	GCCCCGGTTTCAGGCGGTTCATGCAAGCGGCGCGACCCACAGCTCTCTGGGCGGTGCGCTGCACT	120	
QY	154	CTGCTGTTGTTGTTCCGCGGCGCGCGGCTGCGCGGCTGCGCGGCTGCGCGGCGCTGCGCGGCGCTG	213	
DB	121	CTGCTGTTGTTGTTCCGCGGCGCGCGGCTGCGCGGCTGCGCGGCTGCGCGGCGCTGCGCGGCGCTG	180	
QY	214	GTTCCCGTGTGCGCTGCGAGCCGTGCGACGCGCGTGCATGCGCCAGTGCAGCGCTCCG	273	
DB	181	GTTCCCGTGTGCGCTGCGAGCCGTGCGACGCGCTGCATGCGCCAGTGCAGCGCTCCG	240	
QY	274	CCGCGCGTGTGCGGAGCTGTGCGGAGCGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG	333	
DB	241	CCGCGCGTGTGCGGAGCTGTGCGGAGCGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG	300	
QY	334	CTGAGCAGGCGCGAGCGCTGCGGCATCTACACGAGCGCTGTGGCTCCGCGCTTCCGCTGC	393	
DB	301	CTGAGCAGGCGCGAGCGCTGCGGCATCTACACGAGCGCTGTGGCTCCGCGCTTCCGCTGC	360	
QY	394	CAGCGCTCGCCCGACGAGGCGGACCGCTGCGAGCGCTGCTGAGCGCGCGCGGCGCTGC	453	
DB	361	CAGCGCTCGCCCGACGAGGCGGACCGCTGCGAGCGCTGCTGAGCGCGCGCGGCGCTGC	420	
QY	454	GTCAACGCTAGTGCCTGTCAGCGCGCTGCGGCGCTTACCTGCTGCCAGCGCGCGCAGTCCA	513	
DB	421	GTCAACGCTAGTGCCTGTCAGCGCGCTTACCTGCTGCCAGCGCGCGCAGTCCA	480	
QY	514	GGAATGCTAGTGCCTGCGGAGGAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	573	
DB	481	GGAATGCTAGTGCCTGCGGAGGAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	540	
QY	574	TCCAGCAGCGACCGGGTGTCTGATCCCAAGTTCCACCCCGCTCCATTCCAAAGATAATCATC	633	
DB	541	TCCAGCAGCGACCGGGTGTCTGATCCCAAGTTCCACCCCGCTCCATTCCAAAGATAATCATC	600	
QY	634	ATCAAGAAAGGCGATGCTAAAGACAGCGCGCGCTACAAAGTTGCTACGAGTCTCAGAGC	693	
DB	601	ATCAAGAAAGGCGATGCTAAAGACAGCGCGCGCTACAAAGTTGCTACGAGTCTCAGAGC	660	
QY	694	ACAGATACCCAGAACCTTCTCCTCCGAGTCCAAAGCGGAGACAGAAATATGCTCCCTGCCGT	753	
DB	661	ACAGATACCCAGAACCTTCTCCTCCGAGTCCAAAGCGGAGACAGAAATATGCTCCCTGCCGT	720	
QY	754	AGAGAAATGGAAGACACACTGAATCACTGAAATTCCTCAATGTGCTGAGTCCGAGGGT	813	
DB	721	AGAGAAATGGAAGACACACTGAATCACTGAAATTCCTCAATGTGCTGAGTCCGAGGGT	780	
QY	814	GTCACACTTCCCACTGTGACAGAAAGGATTTTATAGAAAAAGACAGTGTGCGCCCTTCC	873	

```

1954 ATATCAGAACAGAGAGAGAGATGGCTTCTGGGAGGCCATCCAGGACACTGGGAGCAC 2013
1960 ATATCAGAACAGAGAGAGAGATGGCTTCTGGGAGGCCATCCAGGACACTGGGAGCAC 1979
1966 ATATCAGAACAGAGAGAGAGATGGCTTCTGGGAGGCCATCCAGGACACTGGGAGCAC 2039
1972 ATATCAGAACAGAGAGAGAGATGGCTTCTGGGAGGCCATCCAGGACACTGGGAGCAC 2073
1978 ATATCAGAACAGAGAGAGAGATGGCTTCTGGGAGGCCATCCAGGACACTGGGAGCAC 2099
2004 ACACATATATGACAGAGAGAGATGGCTTCTGGGAGGCCATCCAGGACACTGGGAGCAC 2133
2010 ACACATATATGACAGAGAGAGATGGCTTCTGGGAGGCCATCCAGGACACTGGGAGCAC 2099
2134 TAGTAAGATCTATATGACAGAGAGATGGCTTCTGGGAGGCCATCCAGGACACTGGGAGCAC 2193
2140 TAGTAAGATCTATATGACAGAGAGATGGCTTCTGGGAGGCCATCCAGGACACTGGGAGCAC 2159
2194 GTGGCCATGACTGAGAGAGAGAGATGGCTTCTGGGAGGCCATCCAGGAGGGGCA 2253
2160 GTGGCCATGACTGAGAGAGAGAGATGGCTTCTGGGAGGCCATCCAGGAGGGGCA 2219
2254 ACCCAAGAGGTCTGGAGAGAGAGATGGCTTCTGGGAGGCCATCCAGGAGGGGCA 2313
2220 AACCAAGAGGTCTGGAGAGAGAGATGGCTTCTGGGAGGCCATCCAGGAGGGGCA 2279
2314 TGTGGACACAGCTGATAGAGAGATGGCTTCTGGGAGGCCATCCAGGAGGGGCA 2373
2280 TGTGGACACAGCTGATAGAGAGATGGCTTCTGGGAGGCCATCCAGGAGGGGCA 2339
2374 TACCTATTAGCTTTCTTTATTTTAACTTTTGGGGGAAAAAGTATTTTGAAGT 2433
2340 TACCTATTAGCTTTCTTTATTTTAACTTTTGGGGGAAAAAGTATTTTGAAGT 2399
2434 TGTCTTGAAGTATTTTAACTTTTAACTTTTGGGGGAAAAAGTATTTTGAAGT 2493
2400 TGTCTTGAAGTATTTTAACTTTTAACTTTTGGGGGAAAAAGTATTTTGAAGT 2459
2494 A 2494
2460 A 2460

```

```

RESULT 7
HUMFIBPA 2474 bp mRNA linear PRI 08-NOV-1994
OCUS Human growth hormone-dependent insulin-like growth factor-binding
DEFINITION protein mRNA, complete cds.
ACCESSION M31159.1 GI:183115
VERSION insulin-like growth factor binding protein.
KEYWORDS insulin-like growth factor (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2474)
AUTHORS Mammalia; Eutheria; Chordata; Vertebrata; Euteleostomi;
Wood, W.J., Cachianes, G., Henzel, W.J., Winslow, G.A., Spencer, S.A.,
Heilmann, R., Martin, J.L., and Baxter, R.C.
TITLE Cloning and expression of the growth hormone-dependent insulin-like
growth factor-binding protein
JOURNAL Mol. Endocrinol. 2 (12), 1176-1185 (1988)
MEDLINE 89112197
PUBMED 2464130
COMMENT Original source text: Human plasma, cdna to mRNA, clone BP-53.
FEATURES
Location/Qualifiers
1..2474
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="7p13-p12"
1..2474
/gene="IGFBP1"
<1..2474
/gene="IGFBP1"

```

```

/product="GFIBP mRNA"
110..985
/gene="IGFBP1"
/note="insulin-like growth factor-binding protein"
/codon_start=1
/protein_id="AA52541.1"
/db_xref="GI:183115"
/translation="MORAPTLMAAALTLVLRGPPVVARAGASSGGLGVVRCPCD
ARALACAPPAVCAELVREPGCCCLTALCEGQPCGIYTERCGSLRCQSPDAR
PLQALLDGRGLCVNAGVRLRAYLLPAPFAGNASEBEDRAGSVSPSVSTHRV
SDKPEHLHSKIIIIKKHAKDSQRYKVDYESQSTDTNFSSKRETEYVPCRRME
DTLNLHKEFLNLSPRGWHIPNCDKGFYKKQCRPSKGRKGRFCMCVXYQQLPGVT
TKKEDVHCYSMOSK"
ORIGIN 132 bp upstream of P1MI site.
Query Match 97.4%; Score 2430.4; DB 9; Length 2474;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2464; Conservative 0; Mismatches 1; Indels 3; Gaps 3;
QY 12 CAGCGGCTTCTGCTGGATTCCACAGCTTCCGCGCGTGTACTGTGCGCCCATCCCTGCG 71
DB 10 CAGCGGCTTCTGCTGGATTCCACAGCTTCCGCGCGTGTACTGTGCGCCCATCCCTGCG 69
QY 72 CGCCACGCTGCCAAGCAGCGTGCCTGGTTGAGGCGTTCATGACGCGCGCGACCCAC 131
DB 70 CGCCACGCTGCCAAGCAGCGTGCCTGGTTGAGGCGTTCATGACGCGCGCGACCCAC 129
QY 132 GCTTGGGGCGCTGCGCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 191
DB 130 GCTTGGGGCGCTGCGCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 189
QY 192 TGGCGGAGCTCGGGGGCTTGGTCCCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 251
DB 190 TGGCGGAGCTCGGGGGCTTGGTCCCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 249
QY 252 ACTGGCCAGTGGCGGCTTCCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 311
DB 250 ACTGGCCAGTGGCGGCTTCCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 309
QY 312 CGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 371
DB 310 CGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 369
QY 372 CTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 431
DB 370 CTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 429
QY 432 GCTGGAGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 491
DB 430 GCTGGAGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 489
QY 492 GCTGCCAGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 551
DB 490 GCTGCCAGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 549
QY 552 CAGTGTGGAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 611
DB 550 CAGTGTGGAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 609
QY 612 CCTCATTCAAAGATATATCATCATCAAGAAAGGCGATGCTAAAGACAGACGCGCGCTACAA 671
DB 610 CCTCATTCAAAGATATATCATCATCAAGAAAGGCGATGCTAAAGACAGACGCGCGCTACAA 669
QY 672 AGTTGACTAGAGTCTCAGAGCAGATACCCAGAGCTTCTCTCCGAGTCCAGGCGGA 731
DB 670 AGTTGACTAGAGTCTCAGAGCAGATACCCAGAGCTTCTCTCCGAGTCCAGGCGGA 729
QY 732 GACAGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 791
DB 730 GACAGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 789
QY 792 CAATGTGCTGAGTCCCGAGGGGTGTACATTTCCCACTGTGACAGAGAGGGATTTTATAA 851

```


Db	2130	GATGAATGTTAGAGTCTATTGATACAACTGTGCCNAGACTGAGGAAGAGCTCA	2189
QY	2221	CGCCAGAGACTGGGTGCTCTCCGAGAGCCAAAACCAAGAGGCTGTGCAAGTCAGG	2280
Db	2190	CGCCAGAGACTGGGTGCTCTCCGAGAGCCAAAACCAAGAGGCTGTGCAAGTCAGG	2249
QY	2281	CTCAGGAGACTCTGCCCTGCTCAGACCTCGGTGTGGACACACGCTGCAATAGAGCTCTC	2340
Db	2250	CTCAGGAGACTCTGCCCTGCTCAGACCTCGGTGTGGACACACGCTGCAATAGAGCTCTC	2309
QY	2341	CTTGAAAAACAGAGGGGTCTCAAGACATCTGCCTCACTTACGCTTTCTTTATTTTTT	2400
Db	2310	CTTGAAAAACAGAGGGGTCTCAAGACATCTGCCTCACTTACGCTTTCTTTATTTTTT	2369
QY	2401	AACCTTTTGGGGGAAAAGTATTTTGGAGAAGTTTGTCTTGCAATGATTTATAAATAGT	2460
Db	2370	AACCTTTTGGGGGAAAAGTATTTTGGAGAAGTTTGTCTTGCAATGATTTATAAATAGT	2429
QY	2461	AAATAAAGTTTTTACCATT	2479
Db	2430	AAATAAAGTTTTTACCATT	2448
RESULT 9			
AK055632			
LOCUS	AK055632	2448 bp mRNA linear	PRI 01-AUG-2002
DEFINITION	Homo sapiens cDNA FLJ13070 fis, clone HSYRA2001332, highly similar to INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 3 PRECURSOR.		
ACCESSION	AK055632		
VERSION	AK055632.1 GI:16550409		
KEYWORDS	oligo capping; fis (full insert sequence).		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 Tashiro,H., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiyama,M., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A., Ohshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S., Naganari,K., Masuho,Y., Nagai,K. and Isogai,I.		
TITLE	NEDO human cDNA sequencing project		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 2448)		
AUTHORS	Isogai,T., Otsuki,T. and Sugiyama,T.		
TITLE	Direct Submission		
JOURNAL	Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0912, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)		
COMMENT	NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5' - 3' -end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: RAB and HRI.		
FEATURES	Location/Qualifiers		
source	1. .2448		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="HSYRA2001332"		
	/cell_type="synovocytes from rheumatoid arthritis (HS-RÅ)"		
	/clone_lib="HSYRA2"		
	/note="cloning vector: pME18SFL3-primary culture, synovocytes from rheumatoid arthritis"		
ORIGIN			

2y 1021 ATGCCTATTATTTGCACAAAAGACTGCCAAGGACATGACCAGCAGCTGGCTACAGCCTCGA 1080
Db 990 ATGCCTATTATTTGCACAAAAGACTGCCAAGGACATGACCAGCAGCTGGCTACAGCCTCGA 1049
2y 1081 TTTATATTTCTGTTTGGTGGAACTGATTTTTTTTTTAAACCAAGTTTGAAGAGGTTT 1140
Db 1050 TTTATATTTCTGTTTGGTGGAACTGATTTTTTTTTTAAACCAAGTTTGAAGAGGTTT 1109
2y 1141 TTGAATGCTATGTTTCTTTGAATGGTAAACTTGAGCATCTTTTCACTTTCCAGTAGT 1200
Db 1110 TTGAATGCTATGTTTCTTTGAATGGTAAACTTGAGCATCTTTTCACTTTCCAGTAGT 1169
2y 1201 CAGCAAGAGCAGTTTGAATTTCTTGTCCGTCCTATCAAAATATTCAGAGACTCGAGC 1260
Db 1170 CAGCAAGAGCAGTTTGAATTTCTTGTCCGTCCTATCAAAATATTCAGAGACTCGAGC 1229
2y 1261 ACAGCACCAGACTTCATGGCCCGTGGAAATGCTCACCATGTTGGTGAAGCGGCCGA 1320
Db 1230 ACAGCACCAGACTTCATGGCCCGTGGAAATGCTCACCATGTTGGTGAAGCGGCCGA 1289
2y 1321 CCAGTGAATTTGTGACTTAGCGGCTGTGTTGCTATGTAGAGAACCGCTTCAACCCCA 1380
Db 1290 CCAGTGAATTTGTGACTTAGCGGCTGTGTTGCTATGTAGAGAACCGCTTCAACCCCA 1349
2y 1381 CTCCTCCGTACAGTCGCGACAGGCTTTATCGAATAGGAAACCTTTAAACCCCGGTCAT 1440
Db 1350 CTCCTCCGTACAGTCGCGACAGGCTTTATCGAATAGGAAACCTTTAAACCCCGGTCAT 1409
2y 1441 CCGGACATCCCAAGCATGCTCTGGAGCTCACAGCCTTCTGTGGTGTCTATTTCTGAAC 1500
Db 1410 CCGGACATCCCAAGCATGCTCTGGAGCTCACAGCCTTCTGTGGTGTCTATTTCTGAAC 1469
2y 1501 AAGGGCTGGAATCCTCAACCAAGAAATGTTTATGTTCTTCAAGTGACCTGACTGCTT 1560
Db 1470 AAGGGCTGGAATCCTCAACCAAGAAATGTTTATGTTCTTCAAGTGACCTGACTGCTT 1529
2y 1561 GGGACATTTGGAGAAATAGTGGAGTCTACTTGTTTTAAAAATATGATATCTAAGAA 1620
Db 1530 GGGACATTTGGAGAAATAGTGGAGTCTACTTGTTTTAAAAATATGATATCTAAGAA 1589
2y 1621 TGTTCTAGGCACTCTGGACCTTATAAGCGAGGTATTTGGGCCCTCTCTTCAGGA 1680
Db 1590 TGTTCTAGGCACTCTGGACCTTATAAGCGAGGTATTTGGGCCCTCTCTTCAGGA 1649
2y 1681 TCCTTCTGAACACATGCCCCAGTCGAAAGGCCAGGATGGCTTTTGTGCGGCCCTGGG 1740
Db 1650 TCCTTCTGAACACATGCCCCAGTCGAAAGGCCAGGATGGCTTTTGTGCGGCCCTGGG 1709
2y 1741 GTAGGGGACAGAGACAGGAGAGTCAAGCTCCACATTCAGAGCATCAAGTAAT 1800
Db 1710 GTAGGGGACAGAGACAGGAGAGTCAAGCTCCACATTCAGAGCATCAAGTAAT 1769
2y 1801 GGCACAATTTCTCGGATGACTGCAGAAAAATAGTGTGTTTGTAGTTTCAACAACTCAAGACGA 1860
Db 1770 GGCACAATTTCTCGGATGACTGCAGAAAAATAGTGTGTTTGTAGTTTCAACAACTCAAGACGA 1829
2y 1861 AGCTTATTTCTGAGGATAAGCTCTTTTAAAGCAAGCTTTATTTTATCTCTCATCTTTT 1920
Db 1830 AGCTTATTTCTGAGGATAAGCTCTTTTAAAGCAAGCTTTATTTTATCTCTCATCTTTT 1889
2y 1921 GTCCTCTTAGCACAATGTAAAAAGAAATAGTAATATCAGAACAGGAGAGGAAATGGCT 1980
Db 1890 GTCCTCTTAGCACAATGTAAAAAGAAATAGTAATATCAGAACAGGAGAGGAAATGGCT 1949
2y 1981 TGCTGGGAGCCCATCCAGGACACTGGGAGCACATAGAGATTCACCATGTTTGTGAAC 2040
Db 1950 TGCTGGGAGCCCATCCAGGACACTGGGAGCACATAGAGATTCACCATGTTTGTGAAC 2009
2y 2041 TTAGAGTCAATCTCATGCTTTCTTTTATATTTACATATATGCAGAGAGATATGTTT 2100
Db 2010 TTAGAGTCAATCTCATGCTTTCTTTTATATTTACATATATGCATAGAGATATGTTT 2069
2y 2101 TTGTTAACTTTGTATACAAACATAGCCCCCAATATAGTAAGATCTTATAGTAATCTCTA 2160

Db 2070 TTGTTAACTTTGTATACAAACATAGCCCCCAATATAGTAAGATCTTATAGTAATCTCTA 2129
2y 2161 GATGAAATGTTTAGAGATGCTATTTGATACAACTGTGGCCATGACTTGAGGAAAGAGCTCA 2220
Db 2130 GATGAAATGTTTAGAGATGCTATTTGATACAACTGTGGCCATGACTTGAGGAAAGAGCTCA 2189
2y 2221 CGCCAGAGACTGGGCTGCTCTCCGGAGGCCAAACCCAGAAAGTCTGCAAAAGTCAGG 2280
Db 2190 CGCCAGAGACTGGGCTGCTCTCCGGAGGCCAAACCCAGAAAGTCTGCAAAAGTCAGG 2249
2y 2281 CTCAGGGAGACTTGCCTGCTGTCAGACCTCGGTGTGGACACACGCTGCATAGAGCTCTC 2340
Db 2250 CTCAGGGAGACTTGCCTGCTGTCAGACCTCGGTGTGGACACACGCTGCATAGAGCTCTC 2309
2y 2341 CTTGAAAAACAGAGGGGTCTCAAGACATTTCTCCCTACCTATTAGCTTTTCTTTATTTT 2400
Db 2310 CTTGAAAAACAGAGGGGTCTCAAGACATTTCTCCCTACCTATTAGCTTTTCTTTATTTT 2369
2y 2401 RACTTTTTTGGGGGAAAGTATTTTGGAGAGTTTGTCTTGAATGTATTTATATAATAGT 2460
Db 2370 RACTTTTTTGGGGGAAAGTATTTTGGAGAGTTTGTCTTGAATGTATTTATATAATAGT 2429
2y 2461 AAATAAAGTTTTTACCATT 2479
Db 2430 AAATAAAGTTTTTACCATT 2448
RESULT 10
AK122873
LOCUS
DEFINITION
Homo sapiens cDNA FLJ16507 fis, clone HCHON2000364, highly similar to INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 3 PRECURSOR.
ACCESSION
AK122873
VERSION
AK122873.1 GI:34528071
KEYWORDS
clog capping; fis (full insert sequence).
SOURCE
Homo sapiens
ORGANISM
Homo sapiens
REFERENCE
1
Tanigami, A., Fujiwara, T., Shibahara, T., Goto, Y., Hirao, M., Ozaki, K., Shimizu, F., Wakabe, H., Ono, T., Hishigaki, H., Watanabe, T., Ozaki, K., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Wagatsuma, M., Murakawa, K., Kaneshiro, K., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Suganac, S., Negahari, K., Masuho, Y., Nagai, K. and Isogai, T.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 2430)
Isogai, T. and Yamamoto, J.
Direct Submission
Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
(E-mail: genomics@nri.co.jp, tel: 81-438-52-3975, fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.
Location/Qualifiers
1..2430
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HCHON2000364"
/cell_type="chondrocytes (HC)"
/clone_lib="HCHON2"

/note="cloning vector: pME18SFL3-primary culture, chondrocytes"

ORIGIN

Query Match 92.6%; Score 2309.4; DB 9; Length 2430;
Best Local Similarity 96.9%; Pred. No. 0;
Matches 2403; Conservative 0; Mismatches 6; Indels 70; Gaps 2;
QY 1 GCGCTGAGGATCAGCGCTTCTGCTCGATTCACAGCTTCGCGCGCTGACTGTGCGC 60
DB 22 GCGCTGAGGATCAGCGCTTCTGCTCGATTCACAGCTTCGCGCGCTGACTGTGCGC 81
QY 61 CCATCCCTGCGCGCCAGCTGCCAAGCAGCGTCCCGGTTGAGCGGTCATCAGCGG 120
DB 82 CCATCCCTGCGCGCCAGCTGCCAAGCAGCGTCCCGGTTGAGCGGTCATCAGCGG 141
QY 121 GCGGACCCAGCTCTGCGCGCTGCGCTGACTGTGCTGCTGCTGCTGCTGCTGCTG 180
DB 142 GCGGACCCAGCTCTGCGCGCTGCGCTGACTGTGCTGCTGCTGCTGCTGCTGCTG 201
QY 181 GTGGGCGGGCTGGCGGAGCTCGGGGGCTGGGTCCGCTGCTGCTGCTGCTGCTG 240
DB 202 GTGGGCGGGCTGGCGGAGCTCGGGGGCTGGGTCCGCTGCTGCTGCTGCTGCTG 249
QY 241 GACGCGCTGCACTGGCCCAAGTGGCGGCTCCGCGCGCTGCTGCGGAGCTGGTGC 300
DB 250 -----CGC 252
QY 301 GAGCGGGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
DB 253 GAGCGGGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 312
QY 361 TACACCGAGCGCTGTGGCTTCGGGCTTCGGCTGCCAGCGCTGCCGCGCGCGGCG 420
DB 313 TACACCGAGCGCTGTGGCTTCGGGCTTCGGCTGCCAGCGCTGCCGCGCGGCGG 372
QY 421 CTGCGAGCGCTGCTGGAAGCGCGCGGGCTGTGGCTCAAGCTGTGCTGCTGCTG 480
DB 373 CTGCGAGCGCTGCTGGAAGCGCGCGGGCTGTGGCTCAAGCTGTGCTGCTGCTG 432
QY 481 CGCGCTACTGCTGCTGCGAGCGCGCGCTGCTGCGGAAATGCTGAGTCCGAGGAG 540
DB 433 CGCGCTACTGCTGCTGCGAGCGCGCGCTGCTGCGGAAATGCTGAGTCCGAGG 492
QY 541 CGCAGCGCGGCGAGTGTGGAGAGCGCGCTCGGCTCCAGCAGCGCGGCTGTGAT 600
DB 493 CGCAGCGCGGCGAGTGTGGAGAGCGCGCTCGGCTCCAGCAGCGCGGCTGTGAT 552
QY 601 AAGTTCCACCCCTCCATTCAGGATATCATCATCAAGAGGCGCATGCTAAGAC 660
DB 553 AAGTTCCACCCCTCCATTCAGGATATCATCATCAAGAGGCGCATGCTAAGAC 612
QY 661 CAGCGCTACAAAGTTGACTACGAGTCTCAGAGCAGATACCCAGAACTTCTCTCC 720
DB 613 CAGCGCTACAAAGTTGACTACGAGTCTCAGAGCAGATACCCAGAACTTCTCTCC 672
QY 721 TCCAGCGGAGACAGATATGGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
DB 673 TCCAGCGGAGACAGATATGGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 732
QY 781 CTGAAGTTCCTCAATGTGCTGAGTCCCGGCTGCTGCTGCTGCTGCTGCTGCTG 840
DB 733 CTGAAGTTCCTCAATGTGCTGAGTCCCGGCTGCTGCTGCTGCTGCTGCTGCTG 792
QY 841 GATTTTATAGAAAAGCAGTGTGCGCTTCCAGGCGAGGCGGCTTCTGCTGG 900
DB 793 GATTTTATAGAAAAGCAGTGTGCGCTTCCAGGCGAGGCGGCTTCTGCTGG 852
QY 901 TGTGTGATAAGTATGGGAGCGCTTCCCGAGGCTACACCAAGGGAAGAGAGCGTG 960
DB 853 TGTGTGATAAGTATGGGAGCGCTTCCCGAGGCTACACCAAGGGAAGAGAGCGTG 912
QY 961 CACTGCTACAGCATGCAGAGCAAGTAGAGCGCTCGCGCAAGGTTAATGTGGAGCTCA 1020

DB 913 CACTGCTACAGCATGCAGAGCAAGTAGAGCGCTCGCGAAGGTTAATGTGGAGCTCA 972
QY 1021 ATGCGCTTATTTTGCACAAAAGACTGCCAAGGACATGACAGCAGCTGGCTACAGCTCG 1080
DB 973 ATGCGCTTATTTTGCACAAAAGACTGCCAAGGACATGACAGCAGCTGGCTACAGCTCG 1032
QY 1081 TTATATTTCTGTTTGTGGTGAAGTATTTTAAACCAAGTTTAAAGAGGTTT 1140
DB 1033 TTATATTTCTGTTTGTGGTGAAGTATTTTAAACCAAGTTTAAAGAGGTTT 1091
QY 1141 TTGAATGCTATGTTTCTTTTGAATGTTAAACTTGCAGCATCTTTTCACTTTCCAGTAGT 1200
DB 1092 TTGAATGCTATGTTTCTTTTGAATGTTAAACTTGCAGCATCTTTTCACTTTCCAGTAGT 1151
QY 1201 CAGCAAGAGAGCTTTGAATTTTCTTGTGCTTCTATCAAAATATTAGAGACTCGAGC 1260
DB 1152 CAGCAAGAGAGCTTTGAATTTTCTTGTGCTTCTATCAAAATATTAGAGACTCGAGC 1211
QY 1261 ACAGCACCAGACTTTCATGCGCGCTGGAACTGCTCACACATGTTGGTCGAAGCGCGCA 1320
DB 1212 ACAGCACCAGACTTTCATGCGCGCTGGAACTGCTCACACATGTTGGTCGAAGCGCGCA 1271
QY 1321 CCACTGACTTTGTGACTTAGGCGCTGTGCTGCTATGTAGAGAACAGCTTTCACCCCA 1380
DB 1272 CCACTGACTTTGTGACTTAGGCGCTGTGCTGCTATGTAGAGAACAGCTTTCACCCCA 1331
QY 1381 CTCGCCGTACGTGCGCAGAGCTTTATCGAATAGGAAACCTTTAAACCCCGCTCAT 1440
DB 1332 CTCGCCGTACGTGCGCAGAGCTTTATCGAATAGGAAACCTTTAAACCCCGCTCAT 1391
QY 1441 CCGGACATCCCAAGCATGCTCTCGAGCTACAGCTTCTGTGTGCTATTTCTGAAC 1500
DB 1392 CCGGACATCCCAAGCATGCTCTCGAGCTACAGCTTCTGTGTGCTATTTCTGAAC 1451
QY 1501 AAGGGCTGCTGCTCCACAGAGAGATGTTTCTTCAAGTGAAGCTGTGCTGCTT 1560
DB 1452 AAGGGCTGCTGCTCCACAGAGAGATGTTTCTTCAAGTGAAGCTGTGCTGCTT 1511
QY 1561 GGGGACTATTGGAGAAATTAAGTGGAGTCTTCTTGTGTTTAAATAATGTATCTAAGAA 1620
DB 1512 GGGGACTATTGGAGAAATTAAGTGGAGTCTTCTTGTGTTTAAATAATGTATCTAAGAA 1571
QY 1621 TGTCTTAGGCACTCTGGAACTTAAAGCAGAGTATTTGCGGCGCTCTCTTCAGGAA 1680
DB 1572 TGTCTTAGGCACTCTGGAACTTAAAGCAGAGTATTTGCGGCGCTCTCTTCAGGAA 1631
QY 1681 TCTTCTGAGACATGCGCCAGTCCGAAGCGCCAGGATGCTTTTGTGCGGCCCGCTGG 1740
DB 1632 TCTTCTGAGACATGCGCCAGTCCGAAGCGCCAGGATGCTTTTGTGCGGCCCGCTGG 1691
QY 1741 GTAGGAGGACAGAGACAGAGGAGTCCAGCTCCACATTCAGAGGCTACAGTAAT 1800
DB 1692 GTAGGAGGACAGAGACAGAGGAGTCCAGCTCCAGCTCCAGGCTACAGTAAT 1751
QY 1801 GGCACAACTTCTCGGATGACTGCAGAAAATAGTGTGTTTGTAGTTCAAACACTCAAGACGA 1860
DB 1752 GGCACAACTTCTCGGATGACTGCAGAAAATAGTGTGTTTGTAGTTCAAACACTCAAGACGA 1811
QY 1861 AGCTTATTTCTGAGGATAAGCTTTTAAAGGCAAGCTTTTATTTTCACTCTCTCACTTTT 1920
DB 1812 AGCTTATTTCTGAGGATAAGCTTTTAAAGGCAAGCTTTTATTTTCACTCTCTCACTTTT 1871
QY 1921 GTCTCTCTTAGCACAATGTAAAAAGAAATAGTAATATCAGAAACAGGAGGAAATGGCT 1980
DB 1872 GTCTCTCTTAGCACAATGTAAAAAGAAATAGTAATATCAGAAACAGGAGGAAATGGCT 1931
QY 1981 TGTGGGAGCCCATCCAGGACACTGGGAGCACAAGAGATTCAACCATGTTTGTGAAC 2040
DB 1932 TGTGGGAGCCCATCCAGGACACTGGGAGCACAAGAGATTCAACCATGTTTGTGAAC 1991
QY 2041 TTAGAGTCACTTCTCATGCTTTTCTTTTATATTCACATATATGCAGAGAAATATGTTTC 2100

2b	1992	TTGAGGTCATTCTCATGCTCTTTCTTTATTAATTCACACATATATGACAGAGAATATGTTTC	2051
2c	2101	TTGTTAAACATTGTATACAAATAGCCCAATATATAGTAAGATCTATCTAGATAATCCCTA	2160
2d	2052	TTGTTAAACATTGTATACAAATAGCCCAATATATAGTAAGATCTATCTAGATAATCCCTA	2111
2e	2161	GATCAAAATGTTAGAGATGCTATTGATACAACTGTGGCCATGACTGAGGAAAGAGAGCTCA	2220
2f	2112	GATGAATGTTTAGAGATGCTATATATATACACTGTGGCCATGACTGAGGAAAGAGAGCTCA	2171
2g	2221	CGCCAGAGACTGGGCTGCTCTCCCGAGGCGCAAAACCCAGAAAGTCTGGCAAGTCAGG	2280
2h	2172	CGCCAGAGACTGGGCTGCTCTCCCGAGGCGCAAAACCCAGAAAGTCTGGCAAGTCAGG	2231
2i	2281	CTCAGGAGAGACTCTGCCCTGCTCGACACCTCGGTGTGGACACACGCTGCTAGAGCTCTC	2340
2j	2232	CTCAGGAGAGACTCTGCCCTGCTCGACACCTCGGTGTGGACACACGCTGCTAGAGCTCTC	2291
2k	2341	CTTGAAACAGAGGGGCTCTCAAGACATCTGCGCTACCTATTAGCTTTCTTTATTTTTTT	2400
2l	2292	CTTGAAACAGAGGGGCTCTCAAGACATCTGCGCTACCTATTAGCTTTCTTTATTTTTTT	2351
2m	2401	AACCTTTTGGGGGAAAGATTTTTTGAGAAGTTGTCTTGCATGTATTTATAAATAGT	2460
2n	2352	AACCTTTTGGGGGAAAGATTTTTTGAGAAGTTGTCTTGCATGTATTTATAAATAGT	2411
2o	2461	AAATAAAGTTTATACCAATT	2479
2p	2412	AAATAAAGTTTATACCAATT	2430
RESULT 11			
LOCUS	AK075483	2412 bp mRNA linear	PRI 03-SEP-2002
DEFINITION	Homo sapiens cDNA PSEC0177 fis, clone OVARC1000287, highly similar to Human growth hormone-dependent insulin-like growth factor-binding protein mRNA.		
ACCESSION	AK075483	1	GI:22761657
VERSION			
KEYWORDS	oligo capping; fis (full insert sequence).		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 Ota, T., Nishikawa, T., Suzuki, Y., Kawai-Hio, Y., Hayashi, K., Ishii, S., Saito, K., Yamamoto, J., Wakamatsu, A., Nagai, T., Nakamura, Y., Sato, H., Nagahara, K., Sugano, S. and Iwagai, T.		
TITLE	HRI human cDNA sequencing project		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 2412)		
AUTHORS	Iwagai, T. and Yamamoto, J.		
TITLE	Direct Submission		
JOURNAL	Submitted (20-MAR-2002) Takao Iwagai, Helix Research Institute, Genomics Laboratory; 1533-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)		
COMMENT	HRI human cDNA sequencing project; cDNA 5' - 3' - end one pass sequencing, clone selection and full insert sequencing; Helix Research Institute (supported by Japan Key Technology Center etc.); cDNA library construction: Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center.		
FEATURES	Location/Qualifiers		
source	1. 2412		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="OVARC1000287"		
	/cissue_type="ovary, tumor tissue"		
	/cissue_lib="OVARC1"		
	/note="cloning vector: pMF18SFL3"		
	239..240		
misc_feature	/note="86 bases segment is present in M31159 and X64875, probably caused by alternative splicing"		

```
QY 1021 ATGCTTAATTTTGCACAAAAGACTGCCAAGGACATGACAGAGCTGGCTACAGGCTCGA 1080
Db 955 ATGCTTAATTTTGCACAAAAGACTGCCAAGGACATGACAGAGCTGGCTACAGGCTCGA 1014
QY 1081 TTTATATTTCTGTTTGGTGAAGTCAATTTTAAACCAAGTTTAAAGAGGTTT 1140
Db 1015 TTTATATTTCTGTTTGGTGAAGTCAATTTTAAACCAAGTTTAAAGAGGTTT 1073
QY 1141 TTGAATGCTTATGCTTTCTTTCAATGTTAAATCTTGAGCATCTTTTCACTTTCCAGTAGT 1200
Db 1074 TTGAATGCTTATGCTTTCTTTCAATGTTAAATCTTGAGCATCTTTTCACTTTCCAGTAGT 1133
QY 1201 CAGCAAGAGACAGTTTGAATTTCTTGTGCTTCTTATCAAAATATTCAGACATCGAGC 1260
Db 1134 CAGCAAGAGACAGTTTGAATTTCTTGTGCTTCTTATCAAAATATTCAGACATCGAGC 1193
QY 1261 ACAGCACCCAGACATTCATGCGCCGCTGGAAATGCTCACCATGTTGGTGGAGCGGCCGA 1320
Db 1194 ACAGCACCCAGACATTCATGCGCCGCTGGAAATGCTCACCATGTTGGTGGAGCGGCCGA 1253
QY 1321 CCACTGACTTTGTGACTTAGCGGCTGTGTGCTATGTAGAGAACACGCTTCAACCCCA 1380
Db 1254 CCACTGACTTTGTGACTTAGCGGCTGTGTGCTATGTAGAGAACACGCTTCAACCCCA 1313
QY 1381 CTCCCGTACAGTGGCGACAGGCTTTATCGAGAAATAGGAAACCTTTAAACCCCGTCAAT 1440
Db 1314 CTCCCGTACAGTGGCGACAGGCTTTATCGAGAAATAGGAAACCTTTAAACCCCGTCAAT 1373
QY 1441 CCGGACATCCCAACGCAATGCTCTGAGGTACAGGCTTCTGTGGTGTCAATTTGAAAC 1500
Db 1374 CCGGACATCCCAACGCAATGCTCTGAGGTACAGGCTTCTGTGGTGTCAATTTGAAAC 1433
QY 1501 AAGGCGTGTGATCCCTCAACCAAGAGAGATGTTATGCTTCAAGTGACCTGTACTGCTT 1560
Db 1434 AAGGCGTGTGATCCCTCAACCAAGAGAGATGTTATGCTTCAAGTGACCTGTACTGCTT 1493
QY 1561 GGGGACTATTTGGGAAATTAAGTGGAGTCTACTTGTTTAAATAATATGATCTAAGAA 1620
Db 1494 GGGGACTATTTGGGAAATTAAGTGGAGTCTACTTGTTTAAATAATATGATCTAAGAA 1553
QY 1621 TGTTCTAGGCACTCTGGGAACCTATTAAGGCGAGTATTTTCGGGCCCTCTCTTCAGGAA 1680
Db 1554 TGTTCTAGGCACTCTGGGAACCTATTAAGGCGAGTATTTTCGGGCCCTCTCTTCAGGAA 1613
QY 1681 TCTTCTCTGAAGACATGCCCCAGTCCGAAGGCCAGGATGGCTTTTGTGGGGCCCGTGGG 1740
Db 1614 TCTTCTCTGAAGACATGCCCCAGTCCGAAGGCCAGGATGGCTTTTGTGGGGCCCGTGGG 1673
QY 1741 GTAGGAGGACAGAGACAGGAGAGTCAAGCTCCACATTCAGAGGATCAAGTAAT 1800
Db 1674 GTAGGAGGACAGAGACAGGAGAGTCAAGCTCCACATTCAGAGGATCAAGTAAT 1733
QY 1801 GGCACAATTTCTCGATGACTCGAGAAATAGTGTGTTTGTAGTTCAACAACTCAAGACGA 1860
Db 1734 GGCACAATTTCTCGATGACTCGAGAAATAGTGTGTTTGTAGTTCAACAACTCAAGACGA 1793
QY 1861 AGCTTATTTCTGAGGATAAGCTTTTAAAGCAAGCTTTATTTTCACTCTCACTTTT 1920
Db 1794 AGCTTATTTCTGAGGATAAGCTTTTAAAGCAAGCTTTATTTTCACTCTCACTTTT 1853
QY 1921 GTCTCTCTTAGCAATGTTAAAGAAATAGTATATCAAGACAGGAGGAGATGGCT 1980
Db 1854 GTCTCTCTTAGCAATGTTAAAGAAATAGTATATCAAGACAGGAGGAGATGGCT 1913
QY 1981 TGTGGGAGCCCATCCAGACACTGGAGACACATAGAGATTCACCCATGTTTGTGAAC 2040
Db 1914 TGTGGGAGCCCATCCAGACACTGGAGACACATAGAGATTCACCCATGTTTGTGAAC 1973
QY 2041 TTAGAGTCATTTCTCATGCTTTTCTTTATATTCACACATATATGCAGAGAGATGTTTC 2100
Db 1974 TTAGAGTCATTTCTCATGCTTTTCTTTATATTCACACATATATGCAGAGAGATGTTTC 2033
```

```
QY 2101 TTGTTAACTTTGTATACAAACATAGCCCCCAATATAGTAGATCTATACTAGATAATCCTA 2160
Db 2034 TTGTTAACTTTGTATACAAACATAGCCCCCAATATAGTAGATCTATACTAGATAATCCTA 2093
QY 2161 GATCAAAATGTTAGAGTCTTATTTGATACAACTGTGGCCCACTGAGGAAAGGAGCTCA 2220
Db 2094 GATCAAAATGTTAGAGTCTTATATGATACAACTGTGGCCCACTGAGGAAAGGAGCTCA 2153
QY 2221 CGCCACAGAGACTGGGCTGTCTCTCCGGAGGCCAAACCCCAAGAGGCTGTGGCAAAGTCAGG 2280
Db 2154 CGCCACAGAGACTGGGCTGTCTCTCCGGAGGCCAAACCCCAAGAGGCTGTGGCAAAGTCAGG 2213
QY 2281 CTCAGGAGACTCTGCCCTGCTGCAGACCTCGGTGGGACACACGCTGCATAGAGCTCTC 2340
Db 2214 CTCAGGAGACTCTGCCCTGCTGCAGACCTCGGTGGGACACACGCTGCATAGAGCTCTC 2273
QY 2341 CTTGAAACACAGAGGGGTCTCAAGACATTTCTGCCTACCTATTAGCTTTCTTTATTTTTT 2400
Db 2274 CTTGAAACACAGAGGGGTCTCAAGACATTTCTGCCTACCTATTAGCTTTCTTTATTTTTT 2333
QY 2401 AACTTTTTGGGGGAAAAAGTATTTTGGAGAGTTTGTCTTGCATGTATTTATAAATAGT 2460
Db 2334 AACTTTTTGGGGGAAAAAGTATTTTGGAGAGTTTGTCTTGCATGTATTTATAAATAGT 2393
QY 2461 AATAAAGTTTTTACCATT 2479
Db 2394 AATAAAGTTTTTACCATT 2412

RESULT 12
AK095907
LOCUS Homo sapiens cDNA FLJ38588 fis, clone HCHON2009795.
DEFINITION AK095907
ACCESSION AK095907.1 GI:21755258
VERSION AK095907.1
KEYWORDS cligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Oshima, A., Takahashi-Fujii, A., Tanase, T., Imose, N., Takeuchi, K.,
Arita, M., Mueashino, K., Yuuki, H., Hara, H., Sugiyama, T., Irie, R.,
Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y.,
Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H.,
Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K.,
Wagatsuma, M., Murakawa, K., Kanehori, K., Sugiyama, A., Kawakami, B.,
Suzuki, Y., Sugano, S., Nagahari, K., Masuho, Y., Nagai, K. and
Isogai, T.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2364)
AUTHORS Isogai, T. and Yamamoto, J.
TITLE Direct Submission
JOURNAL Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Karusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing;
Research Association for Biotechnology (RAB); cDNA library
construction; Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing; RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing; HRI and
RAB; annotation: HRI and RAB.
FEATURES
source
1. 2364
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HCHON2009795"
/cell_type="chondrocytes (HC)"
/clone_lib="HCHON2"
```

/notes="cloning vector: pMEISFL3-primary culture, chondrocytes"

ORIGIN

Query Match 90.2%; Score 2248.6; DB 9; Length 2364;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 2273; Conservative 0; Mismatches 4; Indels 2; Gaps 2;
 201 CTGGGGGGCTTGGGTCCCGTGGTGGCTGCGCTCGAGAGCGGTGGAGCGCGGTGCACTGSCCCA 260
 88 CTCGGCGGCTTGGGTCCCGTGGTGGCTGCGAGCGGTGGAGCGCGGTGCACTGSCCCA 147
 261 GTGGCGCCCTCCGCGCCCGTGGTGGCGGAGCTGGTGGCGAGCGCGGTGGCGGTGGCTG 320
 148 GTGGCGCCCTCCGCGCCCGTGGTGGCGGAGCTGGTGGCGAGCGCGGTGGCGGTGGCTG 207
 321 CTTGAGCTGGCGCACTGAGCGAGGCCAGCGGTGGCGATCTACACGAGCGGTGGCTC 380
 208 CTTGAGCTGGCGCACTGAGCGAGGCCAGCGGTGGCGATCTACACGAGCGGTGGCTC 267
 381 CGGCTTTCGCTGCGAGCGGTGGCGGAGCGCGGTGGCGATCTACACGAGCGGTGGCTC 440
 268 CGGCTTTCGCTGCGAGCGGTGGCGGAGCGCGGTGGCGATCTACACGAGCGGTGGCTC 327
 441 CCGCGGCTCTGCGTCAACGCTAGTGGCTGAGCGCGCTGGCGCTTACCTGCTGCAGC 500
 328 CCGCGGCTCTGCGTCAACGCTAGTGGCTGAGCGCGCTGGCGCTTACCTGCTGCAGC 387
 501 GCGCGAGCTCCAGGAATGCTAGTGGTGGAGGAGCGCGGTGGCGCTTACCTGCTGCAGC 560
 388 GCGCGAGCTCCAGGAATGCTAGTGGTGGAGGAGCGCGGTGGCGCTTACCTGCTGCAGC 447
 561 GAGCGGCTCGCTTCCAGCAGCGACCGGCTGTGTGATCCCAAGTTCACCGGCTTCCATTC 620
 448 GAGCGGCTCGCTTCCAGCAGCGACCGGCTGTGTGATCCCAAGTTCACCGGCTTCCATTC 507
 621 AAGATTAATCATATCAAGAAAGGCGATGCTAAGACAGCGCGGTGACAACTGACTA 680
 508 AAGATTAATCATATCAAGAAAGGCGATGCTAAGACAGCGCGGTGACAACTGACTA 567
 681 CGAGTCTCAGAGCAGATCCAGAACTTCTCTCGAGTCCAAAGCGGAGACAGAA 740
 568 CGAGTCTCAGAGCAGATCCAGAACTTCTCTCGAGTCCAAAGCGGAGACAGAA 627
 741 TGGTCCCTGCGTGAAGAAATGGAAGACACACTGAATCACTGAAGTTCCTCAATGTCT 800
 628 TGGTCCCTGCGTGAAGAAATGGAAGACACACTGAATCACTGAAGTTCCTCAATGTCT 687
 801 GAGTCCAGGGGTGACACATTCCTCACTGACAAAGGATTTTATGAAGAAAGCA 860
 688 GAGTCCAGGGGTGACACATTCCTCACTGACAAAGGATTTTATGAAGAAAGCA 747
 861 GTGTGCGCCCTTCCAAAGGAGGAGCGGGCTTCTGTGGTGGTGAATAGTATGGCA 920
 748 GTGTGCGCCCTTCCAAAGGAGGAGCGGGCTTCTGTGGTGGTGAATAGTATGGCA 807
 921 GCCTTCCAGGCTACACACAGGGGAGGAGCGTGCCTGCTACAGCATGACAG 980
 808 GCCTTCCAGGCTACACACAGGGGAGGAGCGTGCCTGCTACAGCATGACAG 867
 981 CAAGTAGAGCGCTGCGGCAAGGTAAATGGAGGTCAAAATAGCCTTATTTTGCACAAA 1040
 868 CAAGTAGAGCGCTGCGGCAAGGTAAATGGAGGTCAAAATAGCCTTATTTTGCACAAA 927
 1041 GACTGCCAAGACATGACAGAGCTGCTACAGCTCGATTTATTTCTGTGTGGT 1100
 928 GACTGCCAAGACATGACAGAGCTGCTACAGCTCGATTTATTTCT-TTGTGGT 986
 1101 GAACCTGATTTTTTAAACCAAGTTTGAAGAGGTTTTTGAATGCCCTATGCTTCT 1160
 987 GAACCTGA-TTTTTTAAACCAAGTTTGAAGAGGTTTTTGAATGCCCTATGCTTCT 1045
 1161 TTGAATGGTAAACTTGAGCATCTTTTCACTTCCAGTAGTCAGCAAGAGCAGTTTGAAT 1220

DB 1046 TTGAATGGTAAACTTGAGCATCTTTTCACTTCCAGTAGTCAGCAAGAGCAGTTTGAAT 1105
 QY 1221 TTCTTGTGCGCTTCTATCAAAATATTAGAGACTCGAGCAGCAGCCAGACTTTCATGC 1280
 DB 1106 TTCTTGTGCGCTTCTATCAAAATATTAGAGACTCGAGCAGCAGCCAGACTTTCATGC 1165
 QY 1281 GCCCGTGAATGCTCACACATGTTGGTTCGAAGCGCGGACCACTGACTTTGTGACTTAG 1340
 DB 1166 GCCCGTGAATGCTCACACATGTTGGTTCGAAGCGCGGACCACTGACTTTGTGACTTAG 1225
 QY 1341 GCGGCTGTGTTGCCCTATGTAGAGAACACGCTTACCCCACTCCCGGTACAGTGGGACACA 1400
 DB 1226 GCGGCTGTGTTGCCCTATGTAGAGAACACGCTTACCCCACTCCCGGTACAGTGGGACACA 1285
 QY 1401 GGCCTTATCGAGATAGGAACACCTTTAAACCCCGGTATCCGGACATCCCAACGCAATGC 1460
 DB 1286 GGCCTTATCGAGATAGGAACACCTTTAAACCCCGGTATCCGGACATCCCAACGCAATGC 1345
 QY 1461 TCCTGGAGCTCACAGCCTTCTGTGGTGTCAATTTCTGAACAAAGGCGGTGGATCCCTCAAC 1520
 DB 1346 TCCTGGAGCTCACAGCCTTCTGTGGTGTCAATTTCTGAACAAAGGCGGTGGATCCCTCAAC 1405
 QY 1521 CAAGAAGATGTTTATGCTTCAAGTGACCTGTACTGCTGGGACTATTGGAGAAATA 1580
 DB 1406 CAAGAAGATGTTTATGCTTCAAGTGACCTGTACTGCTGGGACTATTGGAGAAATA 1465
 QY 1581 AGTGGAGTCTACTTGTTTAAATAATATGATCTAAGAAATGTTCTAGGGCACTCTGGGA 1640
 DB 1466 AGTGGAGTCTACTTGTTTAAATAATATGATCTAAGAAATGTTCTAGGGCACTCTGGGA 1525
 QY 1641 ACCTATAAAGGAGGATATTGCGGCCCTCTCTTCAAGAAATCTTCTGAAGACATGGCCC 1700
 DB 1526 ACCTATAAAGGAGGATATTGCGGCCCTCTCTTCAAGAAATCTTCTGAAGACATGGCCC 1585
 QY 1701 AGTGAAGCGCCAGGATGGCTTTTGTGCGGCCCTGAGGAGGAGCAGAGAGACA 1760
 DB 1586 AGTGAAGCGCCAGGATGGCTTTTGTGCGGCCCTGAGGAGGAGCAGAGAGACA 1645
 QY 1761 GGGAGAGTCAGCCTCCACATTCAGAGGATCAAGATTAATGGCAATTTCTCGATGAC 1820
 DB 1646 GGGAGAGTCAGCCTCCACATTCAGAGGATCAAGATTAATGGCAATTTCTCGATGAC 1705
 QY 1821 TGCAGAAATAGTGTGTTGTAGTTCAACAACTCAAGAGCAAGCTTATTCTGAGGATAAG 1880
 DB 1706 TGCAGAAATAGTGTGTTGTAGTTCAACAACTCAAGAGCAAGCTTATTCTGAGGATAAG 1765
 QY 1881 CTCTTTAAAGGCAAGCTTTATTTTCACTCTCATCTTTTGTCTCTTCTAGCAATGTA 1940
 DB 1766 CTCTTTAAAGGCAAGCTTTATTTTCACTCTCATCTTTTGTCTCTTCTAGCAATGTA 1825
 QY 1941 AAAAAGATAGTAATATCAGAACAGGAGGAGGATGGCTTGTGGGAGGCCATCCAGG 2000
 DB 1826 AAAAAGATAGTAATATCAGAACAGGAGGAGGATGGCTTGTGGGAGGCCATCCAGG 1885
 QY 2001 ACACCTGGGAGCACATAGAGATTCAACCCTGTTTGTGAACCTTAGAGTCAATTTCTCATGCTT 2060
 DB 1886 ACACCTGGGAGCACATAGAGATTCAACCCTGTTTGTGAACCTTAGAGTCAATTTCTCATGCTT 1945
 QY 2061 TTCTTTTAAATTCACACATATATGAGAGAGATATGTTCTTTGTTAAACATTTGTATACAC 2120
 DB 1946 TTCTTTTAAATTCACACATATATGAGAGAGATATGTTCTTTGTTAAACATTTGTATACAC 2005
 QY 2121 ATAGCCCCAAATATAGTAAGATCTATCTAGATAATCTTAGATAAATGTTTAGAGATGCT 2180
 DB 2006 ATAGCCCCAAATATAGTAAGATCTATCTAGATAATCTTAGATAAATGTTTAGAGATGCT 2065
 QY 2181 ATTGTATACACTGTGGCCATGACTGAGAAAGAGAGCTCACGCCAGAGACTGGGCTGCT 2240
 DB 2066 ATATGATACAACTGTGGCCATGACTGAGAAAGAGAGCTCACGCCAGAGACTGGGCTGCT 2125
 QY 2241 CTCGCCGAGGCCAAACCCCAAGAGGTTGGCAAAAGTCAGGCTCAGGAGACTCTGCCCTG 2300

[illegible]

[illegible]

QY 1321 CCACTGACTTTGTGACTTAGGGGGTGTGTGCTTATGTAGAGAACAGCGTTTCAACCCCA 1380
 Db 1214 CCACTGACTTTGTGACTTAGGGGGTGTGTGCTTATGTAGAGAACAGCGTTTCAACCCCA 1273
 QY 1381 CTCCTCGTACAGTGGCCACAGCGTTTATCGAATAAGGAACCTTTAAACCCCGTTCAT 1440
 Db 1274 CTCCTCGTACAGTGGCCACAGCGTTTATCGAATAAGGAACCTTTAAACCCCGTTCAT 1333
 QY 1441 CCGGACATCCCAACGCGATGCTCTCGAGGCTCACAGCTTCTGTGGTGCATTTCTGAAAC 1500
 Db 1334 CCGGACATCCCAACGCGATGCTCTCGAGGCTCACAGCTTCTGTGGTGCATTTCTGAAAC 1393
 QY 1501 AAGGGGTGGATCCCTCAACCAAGAAAGATGTTTATGCTTCAAGTGACCTGTACTGCTT 1560
 Db 1394 AAGGGGTGGATCCCTCAACCAAGAAAGATGTTTATGCTTCAAGTGACCTGTACTGCTT 1453
 QY 1561 GGGGACTATTGGAGAAATAAGGTGGAGTCTTACTTGTGTTTAAAAAATATGTATCTAAGAA 1620
 Db 1454 GGGGACTATTGGAGAAATAAGGTGGAGTCTTACTTGTGTTTAAAAAATATGTATCTAAGAA 1513
 QY 1621 TGTCTAGGGCACTCTGGGAACCTATAAAGGACAGGTATTTGGGGCCCTCTCTTTCAGGAA 1680
 Db 1514 TGTCTAGGGCACTCTGGGAACCTATAAAGGACAGGTATTTGGGGCCCTCTCTTTCAGGAA 1573
 QY 1681 TCTTCTGAAGACATGCCAGTCGAAGGCCAGGATGCTTTTGTGGGGCCCGGTGGG 1740
 Db 1574 TCTTCTGAAGACATGCCAGTCGAAGGCCAGGATGCTTTTGTGGGGCCCGGTGGG 1633
 QY 1741 GTAGGAGGACAGAGACAGAGGAGTCAAGCTCCACATTCAGAGGACATCAAGTAAT 1800
 Db 1634 GTAGGAGGACAGAGACAGAGGAGTCAAGCTCCACATTCAGAGGACATCAAGTAAT 1693
 QY 1801 GGCACAACTTCTCGGATGCTCGAGAAATAGTGTGTTGTAGTTCAAACACTCAAGACGA 1860
 Db 1694 GGCACAACTTCTCGGATGCTCGAGAAATAGTGTGTTGTAGTTCAAACACTCAAGACGA 1753
 QY 1861 AGCTTATTTCTGAGTAAGCTCTTTAAAGGCAAGCTTATTTTCACTCTCATCTTTT 1920
 Db 1754 AGCTTATTTCTGAGTAAGCTCTTTAAAGGCAAGCTTATTTTCACTCTCATCTTTT 1813
 QY 1921 GTCTCTCTAGCACTAATGTAAGAAAGTAATATCAGACAGGAAGGATGGCT 1980
 Db 1814 GTCTCTCTAGCACTAATGTAAGAAAGTAATATCAGACAGGAAGGATGGCT 1873
 QY 1981 TGCTGGGGAGCCCATCCAGGACACTGGGAGCACATAGAGATTCACCCATGTTTGTGTAAC 2040
 Db 1874 TGCTGGGGAGCCCATCCAGGACACTGGGAGCACATAGAGATTCACCCATGTTTGTGTAAC 1933
 QY 2041 TTAGAGTCACTCTCATGCTTTTCTTTATATTCACACATATATGCAGAGAGATATGTTT 2100
 Db 1934 TTAGAGTCACTCTCATGCTTTTCTTTATATTCACACATATATGCAGAGAGATATGTTT 1993
 QY 2101 TTGTTAAACATTGTATACACATAGCCCCAAATATAGTAAGATCTATCTAGATAATCCTA 2160
 Db 1994 TTGTTAAACATTGTATACACATAGCCCCAAATATAGTAAGATCTATCTAGATAATCCTA 2053
 QY 2161 GATGAATGTTAGAGATGCTATTTGATACAACTGTGGCCATGACTGAGGAAGGCTCA 2220
 Db 2054 GATGAATGTTAGAGATGCTATATGATACAACTGTGGCCATGACTGAGGAAGGCTCA 2113
 QY 2221 CGCCGAGAGCTGGGCTGCTCTCCGGAGGCCAAACCCAAAGAGGCTGGCAAGTCAGG 2280
 Db 2114 CGCCGAGAGCTGGGCTGCTCTCCGGAGGCCAAACCCAAAGAGGCTGGCAAGTCAGG 2173
 QY 2281 CTCAGGAGACTCTGCCCTGTCTGAGACCTCGGTGTGGACACACAGCTGCATAGAGCTTC 2340
 Db 2174 CTCAGGAGACTCTGCCCTGTCTGAGACCTCGGTGTGGACACACAGCTGCATAGAGCTTC 2233
 QY 2341 CTGGAACACAGAGGGGCTCAAGACATCTCGCTTACCTATAGCTTTTCTTTATTTTTT 2400
 Db 2234 CTGGAACACAGAGGGGCTCAAGACATCTCGCTTACCTATAGCTTTTCTTTATTTTTT 2293
 QY 2401 AACTTTTTTGGGGGAAAGATATTTTGGAGAAGTTTGTCTTGGCAATGTATTTTATAATAGT 2460

Db 2294 AACTTTTTTGGGGGAAAGATATTTTGGAGAAGTTTGTCTTGGCAATGTATTTATAATAGT 2353
 QY 2461 AATATAAGTTTTTACCATT 2479
 Db 2354 AATATAAGTTTTTACCATT 2372

Search completed: March 11, 2004, 03:35:45
 Job time : 6369.58 secs